Institut für Organische Chemie und Biochemie
der Technischen Universität München

Investigations on the Nucleotide Binding Domain of KdpB by NMR Spectroscopy

and

Solution Structure of an $\alpha 4\beta 7$ Integrin Antagonist

Melina Haupt

meinen Eltern

Acknowledgements

First and foremost, I wish to express my sincere appreciation to

Prof. Dr. Horst Kessler.

I deeply appreciate his continuous support over the whole period with all its ups and downs. His encouragement to keep on going and reinforce strength when the first double labelled protein sample went down the drain. For the excellent working conditions with first-rate spectrometers. For giving me the opportunity to present my work at various international conferences and meetings. And finally for some wonderful hiking tours with the group, our *AK-Fahrten*, in the Italian Alps.

I would also like to express my gratitude to *Prof. Dr. Karlheinz Altendorf* from the University of Osnabrück for the possibility to gain insights into the fascinating Kdp system and for his support and vivid interest in our cooperation.

This is where I also want to thank *Dr. Marc Bramkamp* and *Brigitte Herkenhoff-Hesselmann* from the University of Osnabrück for furnishing me with protein samples in ridiculously high concentrations and purities for the NMR investigations and for their introduction into biochemistry during my short stay at the Microbiology Department. And Marc in special for many extensive and informative discussions about the Kdp system.

I am very thankful to *Dr. Murray Coles* for the wonderful cooperation in the completion of the KdpBN project and for proof-reading parts of the manuscript.

I would like to acknowledge the excellent work of our administrators *Dr. Rainer Haeßner* and *Alex Frenzel*, who always took a great effort to keep spectrometers and computers running.

I would like to thank Frau *E. Bruckmaier*, Frau *M. Machule* and Frau *B. Diaw* for their professional work in the office and prompt information when it

was time to prolong the contract of employment.

Among my colleagues who all contributed to the special atmosphere of our group I would like to emphasise those from the 32 304 laboratory. Merci à *Dr. Vincent Truffault* pour m'introduire dans les premier pas de l'RMN. Danke an *Michael John*, *Markus Heller* und *Jochen Klages* für die Diskussionen im Labor und am Spektrometer, die Zusammenarbeit und das angenehme Arbeitsklima.

I would also like to thank *Gustav* (aka Dr. Gerd Gemmecker) for generously distributing measurement time at the spectrometers when time was pressing and for some amusing interruptions of the daily laboratory business.

Furthermore, I would like to acknowledge those who contributed in the completion of the written work:
Kai Tetzlaff for the competent aid in setting up the computer and for his persuasiveness that LaTeX was the right program to choose.
Dr. Guy Deutscher for proof-reading large parts of the manuscript.
And not to forget my friends who were not cross that I stayed absent quite often during the final stage of this work.

Finally, I want to thank my family to whom this work is dedicated and who have a considerable part in the completion of this work. I will always be grateful for their support during my studies, their faith and confidence that I would take the right decision and their help and encouragement when science was not going the way it ought to.

Contents

List of Abbreviations

δ	chemical shift
τ_{mix}	mixing time
2D	two dimensional
3D	three dimensional
4D	four dimensional
kdp	operon of Kdp
K_d	dissociation constant
A-domain	actuator or activation domain in P-type ATPases
ABC	ATP binding cassette
ADP	adenosine-5'-diphosphate
AMP	adenosine-5'-monophosphate
AMP-PNP	adenosine 5'-(β,γ-imido)triphosphate
ATP	adenosine-5'-triphosphate
BMRB	biomagnetic research bank
CAM	cell adhesion molecule
COSY	correlation spectroscopy
CSI	chemical shift index
Da	Dalton (g/mol)
DG	distance geometry
DMSO	dimethyl sulfoxide
DSS	4,4-dimethyl-4-silapentane-1-sulfate-4,4-dimethyl-4-silapentane-1-sulfate
EM	electron microscopy

FID	free induction decay
FITC	fluorescein 5-isothiocyanate isomer I
FPLC	fast protein liquid chromatography
FT	Fourier transformation
GTP	guanosine-5'-triphosphate
H2H3	isolated, recombinant actuator domain of KdpB, including deca-histidinyl tag
H4H5	isolated, recombinant catalytic loop of KdpB, including deca-histidinyl tag
HMBC	heteronuclear multiple bond correlation
HMQC	heteronuclear multiple quantum coherence
HSQC	heteronuclear single quantum coherence
IPTG	isopropyl thiogalactoside
ITP	inosine-5'-triphosphate
K_m	Michaelis-Menten constant
KcsA	potassium channel of *Streptomyces lividans* A
kDa	kilo-Dalton = 10^3 g/mol
Kdp	potassium dependent protein
KdpBN	isolated, recombinant nucleotide-binding domain of KdpB, including deca-histidinyl tag
MAdCAM-1	mucosal addressin cell adhesion molecule-1
MD	molecular dynamics
mM	millimolar
MRI	magnetic resonance imaging
N-domain	nucleotide binding domain in P-type ATPases
Ni-NTA	nickel nitrilotriacetic acid
NMR	nuclear magnitic resonance
NOE	nuclear Overhauser effect
NOESY	nuclear Overhauser effect spectroscopy
P-domain	phosphorylation domain in P-type ATPases
P_i	inorganic phosphate

PAGE	polyacrylamid gelelectrophoresis
PASTA	protein assignment by threshold accepting
PCR	polymerase chain reaction
PDB	protein data bank
ppm	parts per million
RMSD	root mean square deviation
ROE	rotating frame nuclear Overhauser effect
ROESY	rotating frame nuclear Overhauser effect spectroscopy
SDS	sodium dodecylsulfate
STD	saturation transfer spectroscopy
T_1	longitudinal / spin-lattice relaxation time
T_2	transversal / spin-spin relaxation time
TG	thapsigargin
Tic	1,2,3,4-tetrahydro-isoquinoline-3-carboxylic acid
TM	transmembrane α-helix
TNP-ATP	2′,3′-O-(2,4,6-trinitrophenyl)-ATP
TOCSY	total correlation spectroscopy
TSI	transition state I
UV	ultra violet
VCAM-1	vascular cell adhesion molecule-1

Chapter

1

Abstract and Scope of the Work

The major part of this work deals with the elucidation of the structure of the nucleotide binding domain of the KdpFABC complex (Chapters 2–7). P-type ATPases are involved in the active transport of ions through biological membranes. They are responsible for the establishment and maintenance of the electrochemical gradient across these membranes. The KdpFABC complex, a P-type ATPase of *Escherichia coli* is a high-affinity K^+ uptake system that operates only when the cell experiences osmotic stress or K^+ limitation. Unlike other P-type ATPases, it is built up of four non-covalently linked subunits; KdpF, KdpA, KdpB and KdpC. The KdpB subunit, is the actual P-type ATPase, but it lost the ability of autonomous ion transport, by capping three transmembrane helices. Instead, it captured an ion channel-like protein, the KdpA subunit, which now assumes ion transport. In the course of this work the solution structure of the nucleotide binding domain of KdpB was accomplished at a backbone RMSD over ordered residues of 0.17 Å. Furthermore, a model of the AMP-PNP binding mode based on intermolecular distance restraints from an isotope-filtered 2D NOESY spectrum was established. The calculated AMP-PNP binding mode shows the purine ring of

the nucleotide to be *clipped* into the binding pocket via a π-π-interaction to F377 on one side and a cation-π-interaction to K395 on the other. This unique binding mode was corroborated by other measurements as NMR titration and saturation transfer difference spectroscopy. This binding mechanism seems to be conserved in all P-type ATPases, except the heavy metal transporting ATPases (type IB), which do not show any alignment in the nucleotide binding domain for the residues phenylalanine and lysine that are critical for the *clipping* mechanism. Due to its low molecular weight, the Kdp-ATPase is currently grouped as a type IA ATPase, which suggests that it is more closely related to the heavy metal transporting ATPases. Thus it can be concluded that it is misgrouped as it has more similarities to type II–IV ATPases. The KdpB N-domain is the smallest and simplest known for a P-type ATPase, and represents a minimal example of this functional unit. No evidence of significant conformational changes was observed within the N-domain upon nucleotide binding, thus ruling out a role for ATP-induced conformational changes in the reaction cycle.

The structure elucidation of *cyclo*(-Tic-Leu-Asp-Thr-Asp-D-pro-) represents the second part of this work (Chapter 8). Some cyclic hexapeptides with the LDT sequence motive have shown to be powerful $\alpha 4\beta 7$ integrin antagonists. An aromatic residue N-terminal to the LDT sequence proved to be advantageous for biological activity. This resulted in the highly active lead compound *cyclo*(-Leu-Asp-Thr-Ala-D-pro-Phe-). Interestingly, replacement of the phenylalanine by the highly restrained mimic Tic, resulted in an equally active molecule. The distance restraints derived from a ROESY experiment, suggest a flip of the peptide bond between Asp[3] and Thr[4], indicating that there might be at least two conformations. The position of the aromatic ring which is critical for binding to the receptor could be identified to be nearly perpendicular to the peptide backbone in both conformers.

Chapter

2

Biochemical Context

2.1 Types of Transport

Living cells are enclosed by a lipid bilayer to separate them from other cells and the extracellular medium. Despite this quasi-impermeable barrier, efflux of solutes has to be established through the membrane to sustain energetic balance and communication with the environment. To maintain this fragile interplay, membranes are interspersed with numerous proteins, such as specific receptors and highly selective transport systems. Plasma membranes usually consist of approximately 50 % protein, membranes enclosing mitochondria and chloroplasts accumulate up to 75 %, whereas nerve cells, which require to be well isolated, contain no more than 18 % protein. Phylogenetic studies examining the relationship of proteins among the kingdoms of life (Animalia, Plantae, Fungi, Protista and Procaryotae) revealed that enzymes involved in metabolism are extremely divers in all living organisms, but the basic transport systems in cells exhibit stunning similarities. Based on these findings it was suggested that transport proteins and enzymes evolved independently of each other as two distinct classes of proteins from

different precursor peptides.[1] In the course of evolution, requirements and thus complexity of the systems increased, following the putative pathway from channel peptides to proteins, secondary transporters and finally to either primary active transporters or group translocators. According to their mode of action (passive or active) the transport systems are grouped either in channels or pumps.

2.1.1 Channels

Channels form a polar pathway through membranes, enabling the transported solutes to follow their electrochemical gradient. Thus, transport through channels is dependent on the concentration of solute and osmotic pressure is the driving force. The maximum rate of flux attained can be as large as 10^7 ions per second, which is almost the speed of free diffusion in the solvent (10^9 molecules per second).[2] The simplest model of a channel is realised in the water pore aquaporin.[3] It provides high selectivity, preventing proton (H_3O^+) flow while maintaining a very high permeation rate for H_2O. 3D structures of aquaporin show that its architecture allows water molecules to pass only in single file.[4,5] To ensure that the internal pH of the cell is maintained, *proton hopping* through the channel has to be avoided. Therefore positively charged residues are situated in the centre of the channel such that they repel H_3O^+. Furthermore, the local electrostatic field generated by the protein switches polarity in the middle of the channel, forcing the passing water molecules to rotate in such a way that their dipole moments are oriented in opposite directions in the upper and the lower halves of the channel. This reorientation prevents the formation of a continuous network of hydrogen-bonded water molecules across the channel, and thus blocks the passage of protons via *proton hopping*. Gating of aquaporin is still a controversial issue in literature. Whilst most aquaporins seem to be only diffusion controlled, which means that there exists only an open state, at least a subset is thought to be gated. Ye *et al.*[6] suggest
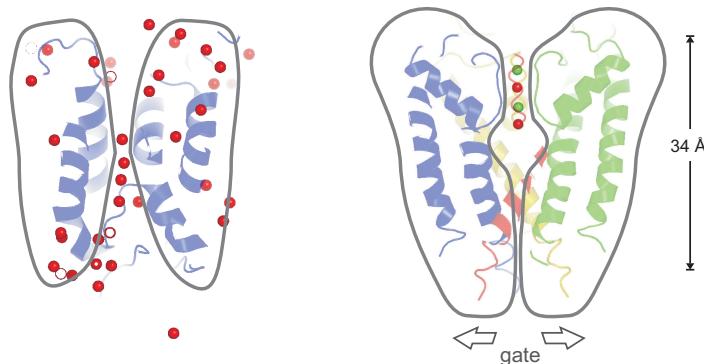
FIGURE 2.1: **Cross section of a water pore and an ion channel.** The crystal structure of aquaporin at 2.2 Å resolution[5] shows that the pore narrows almost symmetrical towards the centre (left panel). In case of the KcsA potassium channel a fourfold selectivity filter controls access to the water filled cavity. A K$^+$ ion and a water molecule occupy alternating binding sites (right panel).

a cohesion-tension model for osmotic gating where tensions within water channels affect the open/closed state by changing the free energy between states. In contrast Tournaire-Roux *et al.*[7] delineate the inhibition of water uptake by anoxia in plants to cytosol acidosis and link aquaporin gating to cytosolic pH. Recent results propose an energy-input model, where mechanical stimuli are responsible for channel constriction provoked by the rate of water flow across channels.[8]

The central concept of ion channels is similar to the one of water channels. High selectivity and fast throughput is achieved by a pore formed of several transmembrane α-helices. Most ion channels consist of four similar subunits of two to up to five transmembrane spanners. Some exist as an assembly in the membrane, others are covalently associated. Gating is a prerequisite and is either ligand or voltage controlled. As early as 1890, Wilhelm Ostwald recognised that electrical currents in living tissues might be caused by ions

moving across cellular membranes.[9] In 1925 the existence of narrow ion channels was proposed.[10] However, it took a few more decades to show that potassium ions move through the membrane in single file.[11]

In the following years the focus was on the ligand-gated acetylcholine receptor. A well studied receptor system stems from the electric ray *Torpedo californica* which led to the first biochemical identification of an ion channel protein. Low-resolution structural studies of the acetylcholine receptor showed a large extracellular funnel leading to a narrow membrane channel.[12,13] The acetylcholine receptor channel which is located at the postsynaptic membrane of nerve cells serves as a popular example for a ligand controlled channel. According to its name it is activated by the hormone acetylcholine. Binding of this neurotransmitter to the postsynaptic membrane abruptly increases the ion permeability. A strong sodium influx and a weaker potassium efflux evoke an action potential and the nerve impulse is transmitted further along the axon. Two molecules of acetylcholine are necessary to open the channel. They are rapidly hydrolysed by an enzyme to acetyl and choline to prevent that the channel stays in an open conformation for longer than one millisecond. Binding of more than two molecules of acetylcholine leads to closure of the channel, the so called desensibilisation.

The access to voltage gated and mechanosensitive ion channels was much more hampered and was only made possible when biochemical knowledge and methods such as cloning, mutagenisation and expression were improved. The first isolation of a sodium channel was obtained from the electric eel *Electrophorus electricus*[14] and suggestions on a putative structure of four domains were made. Sodium channels are abundant in most cells and play a major role in the build-up of the action potential in nerve cells and osmotic regulation. Neurotoxins as abundant in the intestines of blowfish (tetrodotoxine) or snake venom (antigenic peptides and proteins) bind with very high affinity to the sodium channel, rendering it impermeable for Na^+ ions and consequently leading to death by respiratory paralysis. Very low concentrations of tetrodotoxine proved to help tumour patients who do not

respond to common analgetics. Apparently, there are no severe side effects and the venom is currently in clinical phase III trials.[15] Snake venoms are frequently used in treatment of asthmatics because of its muscle relaxing properties.

The first high resolution structure of an ion channel by X-ray crystallography was the one of the potassium channel KcsA (see Figure 2.2). It was solved at a resolution of 3.2 Å in 1998 in MacKinnons laboratory and provided for the first time insights in the construction of the pore at atomic resolution.[16] The KcsA K^+ channel is the prototype of a potassium channel. It is built up of four identical subunits, which create an inverted cone. Each subunit consists of two transmembrane spanning α-helices. A short α-helix between the two transmembrane spanners folds back into the membrane, opening the pore towards the cytoplasm and providing the selectivity filter. In relation to its topology, this subunit is called MPM motive where M stands for a membrane spanning segment and P describes the P-loop (potassium selective motive, it has nothing to do with P-loop ATPases) that folds back into the membrane. The channel as a whole is called consequently 2TM-type (transmembrane span) K^+ channel.

Unlike the water pore, where the selectivity filter is situated in the centre of the channel, in case of KcsA it is placed at the outer end of the pore. It is only 12 Å long, which represents only a third of the total length of the channel (34 Å). The narrow selectivity filter leads into a water-filled cavity in the core of the channel and helix dipoles are positioned so as to overcome electrostatic destabilisation of an ion at the centre of the bilayer. Backbone carbonyl oxygen atoms from the channel protein line the selectivity filter, which is held open by structural constraints to coordinate K^+ ions but not smaller Na^+ ions (see Figure 2.1 and Figure 2.3). The selectivity filter consists of four rings of oxygen atoms and thus four ion binding sites which are alternating filled with water molecules and K^+ ions. When a new potassium ion enters the selectivity filter the K^+ ion next to the extracytoplasmic site moves one binding site deeper into the channel, consequently the first
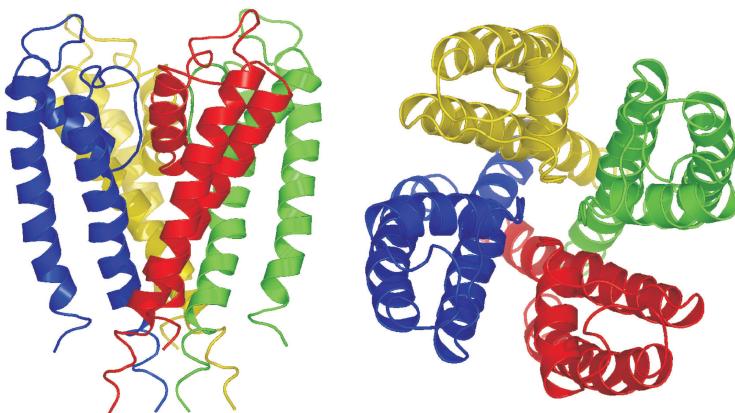
FIGURE 2.2: **Structure of the KcsA potassium channel at 3.2 Å resolution.**[16] The fourfold symmetry of the channel and its stunningly elegant proportions make it a prototype of K[+] channels.

ion is pushed into the water-filled cavity where it is rehydrated to leave the channel by the cytoplasmic mouth.[17]

An exceptional family amongst channel proteins are the so called gap junctions. These channels extend from one cell to another through the extracytoplasm mediating flux between cells. They are encountered in cells that are located in tissues without blood supply, such as in bones or in the lenses of the eye.

Flux of hydrophobic solutes across biological membranes is not a trivial process, considering the need of a fast response to sudden events, such as changes in osmolality, high throughput, high selectivity and manageable regulation. In 2003, the importance of the subject was recognised by awarding the Nobel Prize in Chemistry to Peter Agre for the discovery of aquaporin and to Roderick MacKinnon for elucidation of the structural and mechanistic basis for ion channel function.
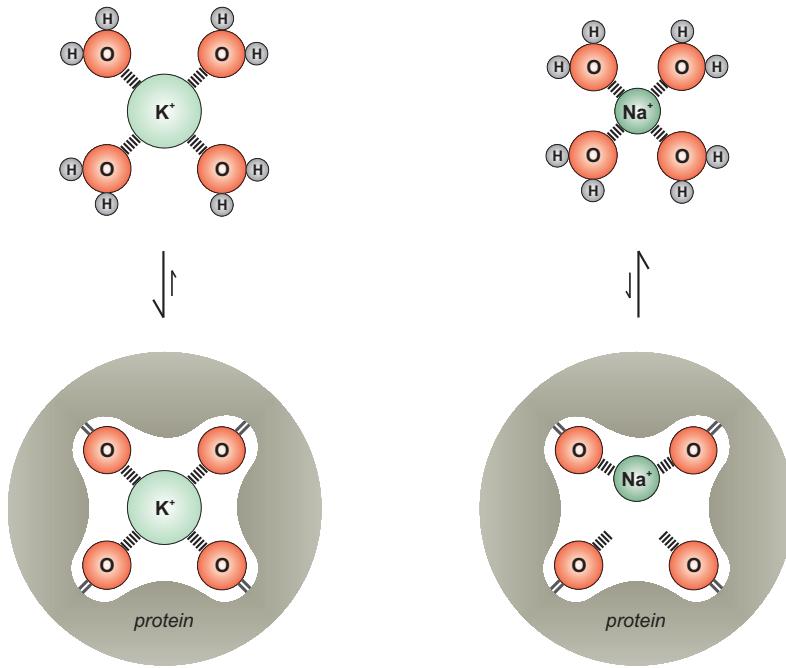
FIGURE 2.3: **Schematic picture of the principle of the selectivity filter in ion channels.** The pore is restricted by four backbone carbonyl oxygen atoms, assuming exactly the same coordination sphere as water molecules do for solvated K^+ ions. When entering the selectivity filter K^+ ions can strip of their hydration coat, the smaller Na^+ ions can't, in consequence they cannot enter the pore. Four layers in series of this type of selectivity filter enhance selectivity even more. For clarity ion radii are not accurately scaled.

2.1.2 Ion Pumps

Potassium is the most abundant solute in all cells, concentrating it against the gradient is vital for life. The concentration of potassium within cells in favour of the more abundant sodium ions is due to its larger radius and thus lower hydration energy. Elevation of the cytoplasmic sodium concentration causes the cell to swell. To maintain a constant volume, sodium has to be removed from the cytoplasm. Exclusion of sodium is also a means of transporting other solutes into the cell using the electrochemical sodium gradient, as it is the case for some of the channels described before. Since Hodgkin and Keynes[11] it is known that upon stimulation of a nerve, sodium ions pour into the cell. To re-establish the membrane potential and regain functionality a counter-transport needs to be started immediately afterwards, restoring the difference in concentration across the membrane. Transport of solutes against their electrochemical gradient requires energy input that is mostly provided by ATP hydrolysis. According to this feature these types of transport proteins were called ion pumps. In 1957 Jens C. Skou from the University of Aarhus, Denmark, isolated an ATP degrading protein from a preparation of crab nerve membranes. This enzyme was dependent on magnesium ions, stimulated by sodium and potassium ions and phosphorylated during the reaction cycle. He could show that enzyme activity was coupled to the action of the ion pump. He had discovered the Na^+,K^+-ATPase, the enzyme responsible for sodium potassium exchange in all living cells.[18] For this discovery Jens C. Skou was awarded the Nobel Prize in Chemistry in 1997 together with Paul D. Boyer and John E. Walker who had pushed ahead the elucidation of the enzymatic mechanism of the ATP synthase.

2.2 P-type ATPases

In the following years many more ATP dependent ion pumps were discovered. Since they all form a phosphorylated intermediate during the course of the reaction they were termed P-type ATPases.[19] Pedersen and Carafoli[20] could show that the active pumping of monovalent and divalent cations, including protons, alkali and heavy metal ions, is mainly accomplished by P-type ATPases. The phylogenetic tree summarises the relationships among the various P-type ATPases (see Figure 2.4).[21] It shows that P-type ATPases are divided into five major branches.

The first main branch are heavy metal pumps (Type IB ATPases). A small subbranch (Type IA) comprises the KdpB proteins of some bacteria, involved in K^+ transport. The second branch is split into several families. Most of them are believed to transport Ca^{2+} (Type IIA and Type IIB). The Na^+,K^+- and H^+,K^+-ATPase are closely related and are designated Type IIC ATPases. A small family of fungal ATPases of unknown function form a subbranch (Type IID ATPases). The third branch covers plasma membrane H^+-ATPases (Type IIIA ATPases) and Mg^{2+}-ATPases from bacteria. The fourth branch assembles an only recently discovered new family of phospholipid pumps (Type IV ATPases). The fifth branch (Type V ATPases) gathers a bunch of pumps having no assigned substrate specificity yet.

Although the sequence homology between the different members of the P-type ATPase family is not very high, some key motives are highly conserved,[22] which suggest a similar modular design for all family members. Common to all P-type ATPases is the conserved DKTGT motive, comprising the invariant aspartic acid residue, which gets phosphorylated during the reaction cycle. Atomic resolution structures by X-ray crystallography at a resolution of 2.6 Å of the calcium pump from the sarcoplasmatic reticulum (SERCA1a) revealed the architecture of the family in detail.[23] The cytoplasmic loops form three separate modules, commonly named A-, P- and N-domains (Fig. 2.5). The P-domain assumes a typical Rossmann-fold
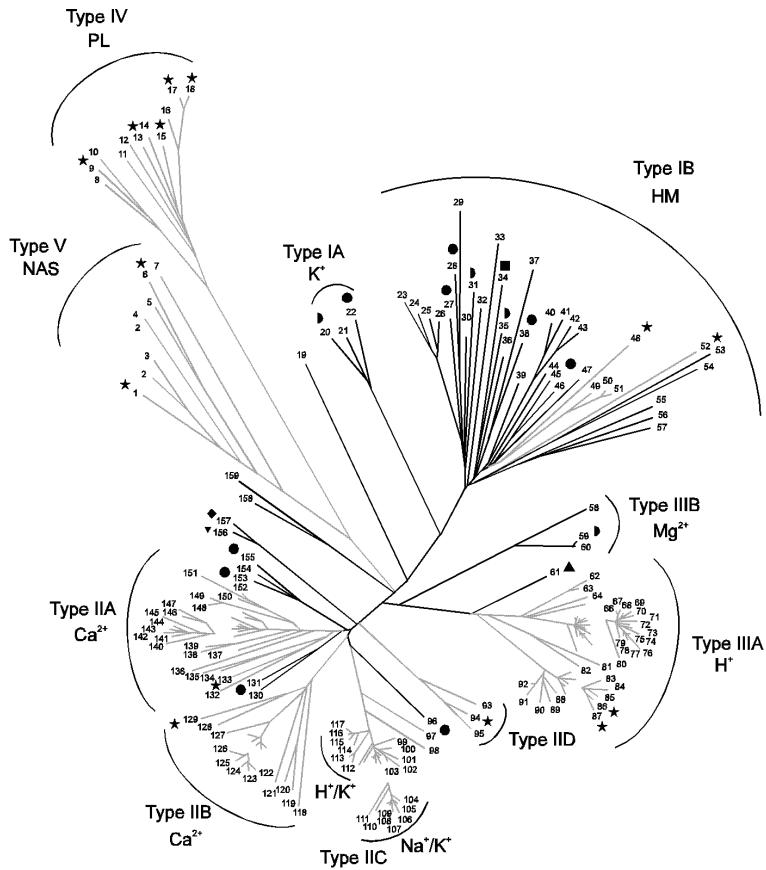
FIGURE 2.4: **Phylogenetic tree of P-type ATPases.** The K$^+$ transporting subunit KdpB of the P-type ATPase Kdp from *E. coli* belongs to type IA, whereas other alkali transporting ATPases are classified in types II–IV. Heavy metal transporting ATPases are of similar size as KdpB and are grouped in type IB. Figure taken from Axelsen & Palmgren.[21]
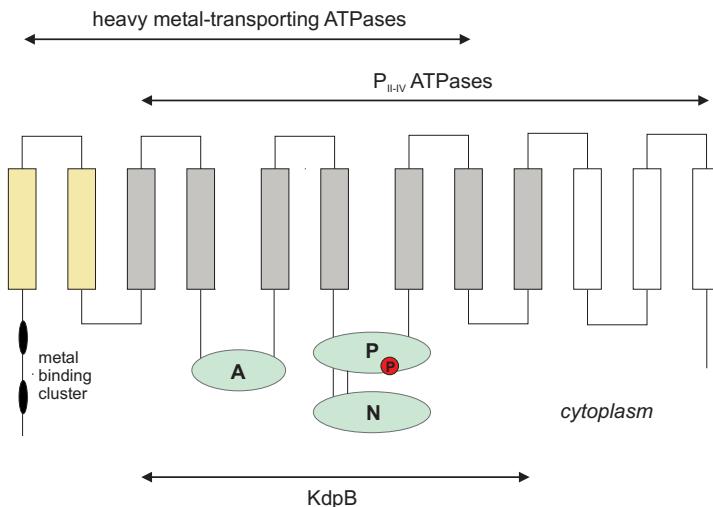
FIGURE 2.5: **Schematic topology of P-type ATPases depicting the modular design.**
The proposed topology of KdpB (type IA) is shown in grey. Heavy metal-transporting
ATPases (type IB) have an N-terminal extension (shown in yellow), whilst P-type
ATPases of the type II-IV have 10 transmembrane helices (extension shown in white).
The phosphorylation site of P-type ATPases is a conserved aspartate residue in the
large cytoplasmic loop (red circle).

and contains the invariant aspartic acid residue that becomes phosphory-
lated during the reaction cycle. The N-domain, comprising the nucleotide
binding site, is inserted into the P-domain, while the A-domain, formed by
another smaller cytoplasmic loop, mediates dephosphorylation of the en-
zyme. Interestingly, the structure reveals no channel-like pathway from the
cytoplasmic side of the membrane to the lumenal side of the membrane.
The high-affinity binding sites for Ca^{2+} are located almost in the centre of
the transmembrane region of the ATPase. Conserved positively charged re-
sidues in the transmembrane α-helix 4 seem to mediate ion translocation
when undergoing a conformational change upon phosphorylation.

2.3 The Kdp System

The Kdp system (potassium dependent protein) is a high-affinity (K_M ~2 mM) K^+ transport system that was originally found in *Escherichia coli*. The membrane-bound KdpFABC complex is an emergency potassium up-take system that is only expressed when the cell suffers osmotic stress, such as potassium limitation or high extracellular osmolality and when other transport systems (TrkG/H, Kup, KtrAB) are unable to meet the cell's need for K^+ ions. This is usually the case at concentrations below 0.1 mM. Nevertheless, the expression of the KdpFABC complex is not only influenced by extracellular K^+ ion concentration but rather by the effect of cytoplasmic K^+ in maintaining the level of turgor needed for growth of bacteria.[24] The expression of the KdpFABC complex is controlled by the regulatory proteins KdpD and KdpE. The KdpFABC complex is rapidly degraded when no longer required. The *kdp* operon also occurs in many other Bacteria. The Genbank Data Base provides some organisms which contain the *kdpFABC* operon such as some Archaea, though without the regulatory genes *kdpD* and *kdpE*.[25] So far it could not be shown to exist in Eukarya, which is plausible considering its function. Bacteria and Archaea, consisting of a single cell, are much more vulnerable with respect to changes in osmolality of their surrounding medium than Eukarya are. Therefore they require a much more elaborate system to adapt quickly to environmental challenges. The KdpFABC complex is unique amongst P-type ATPases in both its function and subunit organisation (see Figure 2.6). Unlike other members of the family, uptake of K^+ ions is not accompanied by the counter-transport of another ion. The unique function of the KdpFABC complex is reflected in an equally unique architecture. Generally, P-type ATPases consist of a central catalytic subunit that facilitates both ion transport and ATP hydrolysis. However, in the KdpFABC complex phosphorylation and ATP hydrolysis are performed in subunit KdpB, whereas K^+ binding and transport have been associated with the KdpA subunit.[24]
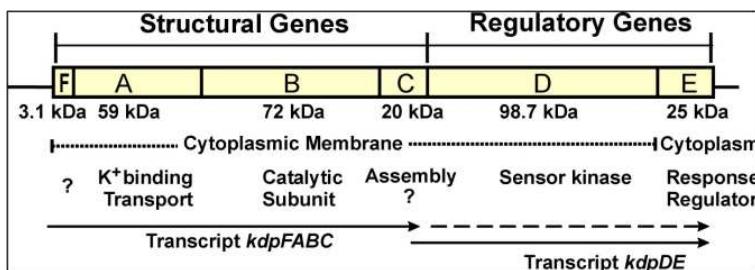
FIGURE 2.6: **The *kdp* operon from *Escherichia coli*.**[24] The structural genes *F*, *A*, *B* and *C* are situated upstream from the regulatory genes *D* and *E* (Marc Bramkamp, personal communication).

2.3.1 The KdpA Subunit

The 59 kDa KdpA subunit is a highly hydrophobic membrane bound protein that spans the membrane 10 times[26] and it has only small hydrophilic extramembranous loops. KdpA mutants exhibiting altered K_m values lend support to the notion that KdpA is the subunit responsible for K$^+$ binding and transport.[24] KdpA is evolutionarily derived from the superfamily of 2TM-type K$^+$ channels (KcsA) in bacteria. The eight middle transmembrane spans are formed by four single-MPM motive subunits. The first and the 10th transmembrane span of KdpA are later evolutionarily achievements. Computational methods based on sequence alignment, taking into account the typical M1-P-M2 segments, could show that the KdpA sequences from 17 different organisms have a pattern of residue conservation consistent with the crystal structure of the KcsA K$^+$ channel.[27] Interestingly, the most highly conserved residues between the families are the central glycines of the P-loop segments, those were previously identified as crucial for K$^+$ selectivity. According to the model of Durell *et al.*[27] the four M2 α-helices form the inner core of the channel, widening it from the inner half of the membrane towards the cytoplasm. The four M1 α-helices are arranged on

the outer surface of the channel protein, where they are highly exposed to the alkyl chains of the lipids. This is reflected in an increased occurrence of lipophilic residues – approximately a third of M1 consists of leucine. The four P-loop segments are orientated with fourfold symmetry around the axis of the transmembrane pore. The first portion of each P segment (P1) assumes an α-helical conformation pointing directly towards the water-filled cavity in the centre of the pore. The latter portion of each P-loop segment (P2) is in a relatively extended conformation, which returns the peptide chain to the outer surface of the membrane and forms the narrowest region of the channel. The K^+ binding sites are formed by the backbone carbonyl oxygen atoms of the four P2 segments, which are oriented toward the axis of the pore. This new model of the KdpA subunit suggest that it could function as a monomer, whereas previous studies assumed a dimeric model of the KdpFABC complex.[24, 26] In contrast to other P-type ATPases and K^+ translocating proteins, the KdpFABC complex has an increased selectivity for K^+ cations.[28] This might be due to the fact that channel proteins, owing to their rigid architecture, show in general a higher ion selectivity as transport proteins.

2.3.2 The KdpB Subunit

The largest subunit KdpB (72 kDa) represents the catalytic entity and shares extensive regions of homology with other P-type ATPases. These motives of homology were first described by Serrano,[22] they are used as a screening tool for this enzyme family within new sequence data. One of the most characteristic motives is the already mentioned DKTGT sequence. It could be shown for KdpB that mutation of aspartate 307, which is the corresponding residue within the DKTGT motive, to asparagine completely abolishes the enzymatic function.[29] Even though its proposed topology matches the key features of all P-type ATPases[19, 28, 30] and is compatible with the headpiece-stalk model, derived from cryo EM pictures for this class of enzymes,[22, 31]

KdpB is much smaller than other P-type ATPases. Hydrophobicity plots suggested the existence of six or seven transmembrane helices[32] instead of ten in case of type II-IV, or eight in case of the heavy metal transporting type IA P-type ATPases.

When the first high resolution crystal structure of a P-type ATPase was published this signified a major breakthrough in the field of ion pump research. Toyoshimas structure of the Ca^{2+}-ATPase of sarcoplasmatic reticulum (SERCA pump) at 2.6 Å resolution[23] in the Ca^{2+}-bound state (E1) clearly shows the modular design of this class of enzymes. Three well separated cytoplasmic domains, commonly termed A-, P- and N-domain, shape the protein. The A-domain comprising the conserved TGES motive is the actuator domain, mediating dephosphorylation of the enzyme. The A-domain is built up of the cytoplasmic loop connecting transmembrane helix 2 and 3 (H2H3 loop) and of the N-terminus of the protein protruding into the cytosol. The β-sheet core is surrounded by four rather short α-helices, a β-hairpin protrudes into the solvent, pointing approximately 90° away from the P-domain. The tip of this β-hairpin comprises the TGES sequence. Mutational studies, altering the TGES motive resulted in an irreversibly phosphorylated enzyme.[33] The largest cytoplasmic loop between transmembrane helix 4 and 5 contains both the P- and the N-domain. The phosphorylation domain (P-domain) assumes a typical Rossmann-fold with an altering pattern of β-strands and α-helices. The nucleotide binding domain (N-domain) is inserted into the P-domain, closely behind the phosphorylation site (D351, in case of SERCA1a). The N-domain of the Ca^{2+}-ATPase is a twisted 9-stranded antiparallel β-sheet surrounded by approximately 7 α-helices and a β-hairpin. The nucleotide binding motive is not as highly conserved as the other characteristic motives of P-type ATPases. Type II-IV P-type ATPases commonly show the KGAPE/D sequence, KdpB supports the KGSVD motive whereas the heavy metal transporting P-type ATPases (type IB) hardly show any alignment in this region (Figure 2.7). Whereas the A- and the P-domain in all P-type ATPases share a similar size and thus
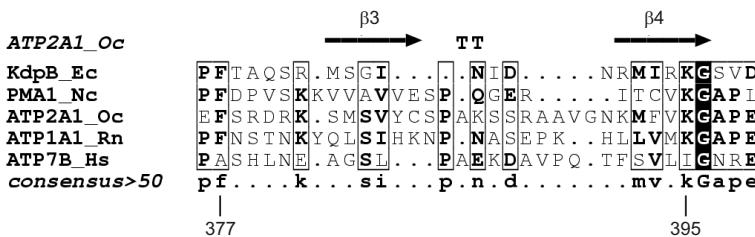
FIGURE 2.7: **Alignment of P-type ATPases at the N-domain specific KGXXD/E motive.** Chosen were ATPases representing different classes according to Axelsen & Palmgren.[21] The abbreviations are Kdp_Ec for KdpB from *Escherichia coli*, PMA1_Nc for H⁺-ATPase from *Neurospora crassa*, ATP2A1:Oc for Ca²⁺-ATPase of *Oryctolagus cuniculus*, ATP1A1_Rn for Na⁺,K⁺-ATPase of *Rattus norvegicus* and ATP7B_Hs for Cu²⁺-ATPase of *Homo sapiens*. Numbering of β-strands and residues was done according to KdpB. For details see ref [34].

most likely a similar fold, the N-domains differ drastically in size.

Homology modelling of KdpB based on the crystal structure of the Ca²⁺-ATPase shows a good agreement in most parts of the molecule, although the overall identity of the core sequences is less than 25 %.[21] The modeling also revealed that KdpB indeed has seven transmembrane spanning segments, instead of the six predicted ones. Since the C-terminal part of the enzyme, where the additional helix happens to be, is mainly hydrophobic, it is difficult to tell from the hydrophobicity plots whether there are two or three transmembrane helices. Another remarkable result of the computer-based modeling is that the nucleotide binding domain of KdpB which is approxiamately 100 amino acids shorter than the N-domain of the SERCA pump showed no structural alignment at all (Figure 2.7).

2.3.3 The Subunits KdpC and KdpF

The 20 kDa KdpC subunit is predicted to have only one membrane spanning α-helix close to its N-terminus. The rest of the protein is exposed to the

cytoplasm. KdpC is absolutely essential for the function of the KdpFABC complex,[24] but its role is still a matter of discussion.[35]

The 3 kDa KdpF peptide is very hydrophobic and comprises only one membrane-spanning segment. Analysis of a *kdpF* deletion strain indicates that the peptide is not essential in vivo, but in vitro the purified KdpFABC complex of such a deletion strain does not exhibit ATPase activity.[36] However, it was possible to recover ATPase activity by the addition of purified KdpF, leading to the assumption that KdpF stabilizes the KdpFABC complex at least in vitro.[37]

Neither KdpC nor KdpF have regions of homology to other proteins within the Genbank Data Base worth mentioning.

2.3.4 Regulation of the kdp Operon - The Subunits KdpD and KdpE

The expression of the *kdpFABC* operon is regulated via KdpD (99 kDa) and KdpE (25 kDa), members of the large family of sensor kinase / response regulator proteins. The membrane bound KdpD serves as an osmosensor. The sensor kinase becomes autophosphorylated when activated and transfers the phosphate group to the smaller, soluble response regulator KdpE in the cytoplasm. The sensor kinase KdpD is anchored to the membrane by four closely-spaced transmembrane spans situated approximately in the middle of the primary sequence. The N- and the C-termini form two large cytoplasmic domains.[38] The C-terminus is homologous to other sensor kinases and hosts the histidine residue (His673) which is the site of phosphorylation.[39] The N-terminal domain shows no similarities to any other proteins of this class, but it was shown to play a crucial role in sensing the signal and modulating the phosphorylation and phosphotransfer activity together with the transmembrane spans.[40] Autophosphorylation of KdpD goes in hand with a conformational change. The trigger for autophosphorylation of KdpD is presumably a combination of sensing the turgor pres-

sure on one side and maintaining a defined pool of cytoplasmic K^+ ions or the rate of K^+ ion transport on the other. A turgor control model has been proposed where the kinase activity is linked to the conformation and hence on the membrane stretch.[24, 41] At normal turgor the conformation of KdpD is such that access to His673 is restricted. Changes of osmolality leading to a decrease of turgor and a falling stretch of the membrane result in a conformational change of KdpD and immediate autophosphorylation by cytoplasmic ATP. In vitro phosphotransfer between KdpD and KdpE takes place rapidly. KdpE bears resemblance to other response regulator proteins.[42] The phosphoryl group is most likely transferred to Asp52, which is conserved among this class of proteins. Phosphorylation of KdpE results in a more than 10-fold enhanced binding ability to DNA, which occurs at the upstream sequence of the *kdpFABC* promoter in the region between -72 and -50 (transcription starts at +1). Phospho-KdpE does not only efficiently stimulate the transcription of the *kdpFABC* operon to RNA, but it also leads by readthrough to additional expression of *kdpD* and *kdpE*, which stimulates the expression even further. To terminate the expression of the KdpFABC complex, phospho-KdpE has to be dephosphorylated. It is suggested that the sensor kinase KdpD has, alike described for other systems (EnvZ/OmpR), phosphatase activity when turgor and membrane stretch have regained a normal level. Consequently, the level of phospho-KdpE and thus transcription of the *kdpFABC* operon could be modulated in dependence on the stimulus, enabling the Kdp system to respond quickly to environmental stimuli.

2.4 The Proposed Reaction Cycle – The E1E2 Model

Quite early Skou and others suggested that the active pumping of ions goes in hand with major changes in the form of the corresponding enzyme.[43] Two main states were proven to exist: The unphosphorylated E1 state with high affinity ion binding sites on the extracytoplasmic site of the membrane
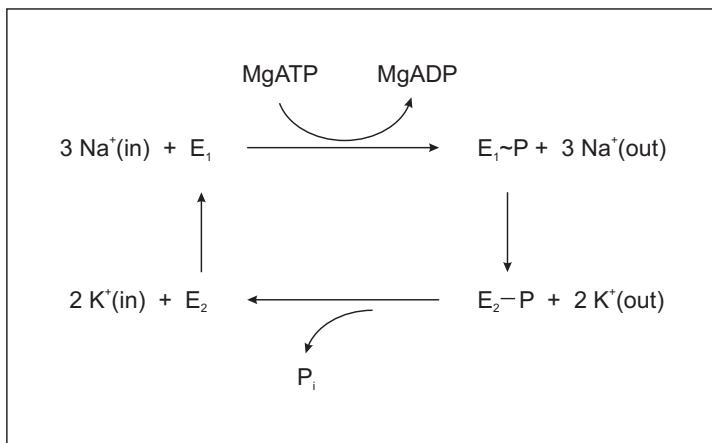
FIGURE 2.8: **The E1E2 reaction cycle of the Na^+,K^+-ATPase according to R. L. Post and R. W. Albers.** Three Na^+ ions are pumped from the cytoplasm to the extracellular side in exchange for two K^+ ions per ATP hydrolysed.

and the phosphorylated E2 state with low affinity ion binding sites on the cytoplasmic side.

In 1967 Albers outlined the biochemical aspects of ion transport and concluded with focus on the Na^+,K^+-ATPase that additional steps have to be implied to explain the behaviour of the enzyme at different experimental conditions.[44] His idea was further elaborated by Post and coworkers[45] and became known as the Post-Albers scheme, shown in Figure 2.8. The reaction cycle starts with a complex of the enzyme with intracellular ATP, Mg^{2+}, and three Na^+ ions. After phosphorylation of the enzyme and release of ADP to the intracellular solution, the three bound Na^+ ions are exchanged for two K^+ ions from the extracellular solution. Eventually inorganic phosphate (P_i) and the two bound K^+ ions are released to the intracellular solution.

A popular interpretation of the mechanism which is still often cited in literature was introduced by DeMeis and Vianna (Figure 2.9).[46] They applied
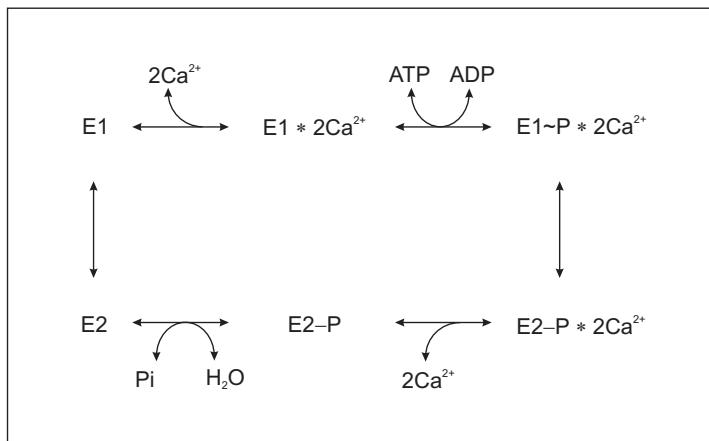
FIGURE 2.9: **The E1E2 reaction cycle of the Ca^{2+}-ATPase of sarcoplasmic reticulum according to DeMeis and Vianna.**[46] This traditional model suggests an spontaneous E2 / E1 transition prior to the binding of two Ca^{2+} ions. The arrows signify that the reaction cycle is reversible under definite circumstances.

the Post-Albers scheme on the Ca^{2+}-ATPase of sarcoplasmic reticulum by maintaining the idea of only two distinct conformational states throughout the whole cycle, but extended it by two steps. Both newly introduced steps are rather doubtful as they deliberate the binding and release of ions from conformational changes. This mechanism implies a spontaneous transition from the E2 to the E1 state prior to ion binding and spontaneous release of the two bound Ca^{2+} ions without any change in conformation. The E1 state then has a micromolar ion binding affinity on the cytoplasmic side for Ca^{2+} ions. The E2 state exhibits low ion binding affinity towards the lumenal side of the membrane. The cycle is reversible, thus the ATPase could in principle bind inorganic phosphate to the E2 state and synthesise ATP by pumping Ca^{2+} ions in the opposite direction.

Both reaction schemes suffer enormous drawbacks as they are restricted

to two conformations, E1 and E2. Furthermore they do not take into account that one main point of enzymatic reactions is the stabilisation of transition states.[47,48]

The missing piece in understanding the mechanism was and is still the mode by which ATP hydrolysis and ion transport could be linked, as they are performed by two distinct parts of the enzyme which are approximately 50 Å away from each other. Electron density maps of two dimensional crystals brought the first rough idea of the structural shape. Stokes and coworkers were able to determine the structure of the Ca^{2+}-ATPase at 8 Å resolution in the presence of *ortho*-vanadate[49] and Kühlbrand and coworkers resolved the structure of the *Neurospora crassa* proton pump at the same resolution without any additives.[50] Interestingly, the cytoplasmic domains of both enzymes were arranged differently, giving support to the idea that the *ortho*-vanadate bound Ca^{2+}-ATPase is in an E2-like state, whereas the H^+-ATPase exhibits the open E1 conformation. The meanings of these conformational changes and their mechanistic roles remained unclear until the first high resolution structure of an P-type ATPase was solved.[23]

2.5 From the Structure Towards a New Understanding of the Reaction Cycle

Previously, biochemical assays had shown that in case of the Na^+,K^+-ATPase the controlled proteolysis with trypsin yields different proteolytic patterns, depending on the presence of either Na^+ or K^+ ions.[51] Furthermore it had been demonstrated that both the Na^+,K^+-ATPase and the Ca^{2+}-ATPase are labelled with the ATP mimic FITC within their active cytoplasmic site, allowing the use of these labelled proteins for numerous investigations on the induced conformational changes.[52]

The high resolution crystal structures of Toyoshima and coworkers of the Ca^{2+}-ATPase in the Ca^{2+}-bound form, E1·Ca^{2+},[23] and thapsigargin-bound

form E2(TG)[53] brought new light into the discussion of ion pumping across
biological membranes (Figure 2.10). They revealed that considerable rear-
rangements of the transmembrane helices take place accompanying Ca^{2+}
dissociation and binding and that they are mechanically linked with equally
large movements of the cytoplasmic domains. Four transmembrane helices
(TM4, 5, 6 and 8) were identified to accomplish the binding of two Ca^{2+} ions.
TM8 does not move during $E1 \cdot Ca^{2+} \rightarrow$ E2(TG) transition, but the others ei-
ther move considerably (TM4), are bent (TM5) or rotated (TM6). As a result
of these intramembrane rearrangements the number of oxygen atoms that
can coordinate to Ca^{2+} decreases. Subsequently Ca^{2+} is released into the
lumen. Hence, the movement of TM4 enables the release of Ca^{2+} and the
binding of the proton which gets transferred to the cytoplasmic side of the
membrane. Transmembrane helix TM5 plays a central role in ion transloca-
tion as it starts in the P-domain in close proximity to the phosphorylation
motive, extends into the stalk segments and finally spans the membrane
without interruption. Bending of it allows different residues on different
faces to be used for cation binding (e.g. in the Na^+,K^+-ATPase serine can be
used instead of glutamate in binding of the larger K^+ ion instead of Na^+).
This kind of switching specific binding site affinity avoids competition be-
tween the binding cations and thus confusion concerning the direction the
corresponding cation is to be moved.

The detailed knowledge of the structure of the Ca^{2+}-ATPase made it now
possible to model the structures of other P-type ATPases accordingly from
less precise data derived from cryo electron microscopy. Within a short
period the three dimensional structure of the Na^+,K^+-ATPase was solved
from electron density maps at 11 Å resolution[54] and at 9.5 Å resolution.[55]
Kühlbrandt and coworkers fitted the data from the Ca^{2+}-ATPase to an 8 Å
map of the proton pump from *Neurospora crassa*.[56] The H^+-ATPase exists
in an inactive and an active form. The crystal structure was obtained on the
inactive form. Activation is achieved by reversible phosphorylation of the
C-terminus (R-domain). Based on their results, they made suggestions on
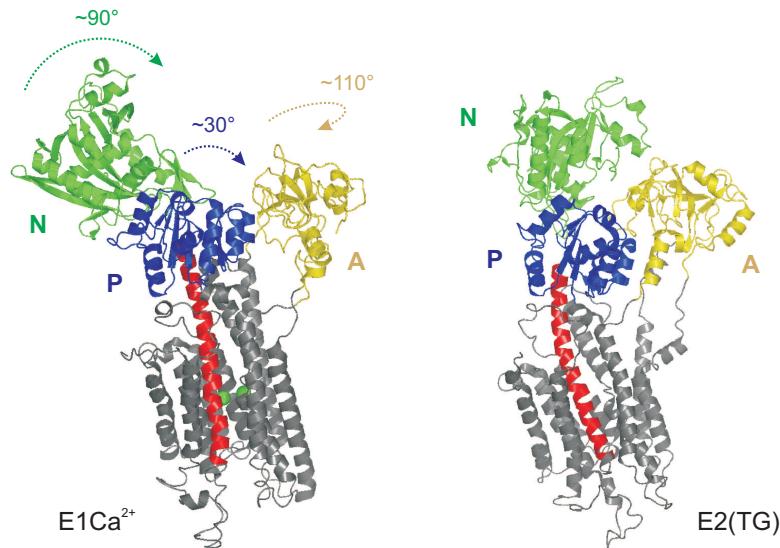
FIGURE 2.10: **The Ca²⁺-ATPase in the "open", Ca²⁺-bound state E1 (1EUL)[23] and in the "closed" thapsigargin-bound state E2 (1IWO).[53]** Arrows indicate the large rigid body movements of the cytoplasmic domains occurring during the E1·Ca²⁺→ E2(TG) transition. Helix 5 was coloured red.

the putative reaction cycle and regulation of the pump. The inactive form showed a high degree of alignment to the closed form of the Ca²⁺-ATPase. Thus, they concluded that the regulatory domain (R) might hold the pump in a closed state and by that prevent the binding of ATP. The prospect of finding a plausible structure functionality relationship has risen a vivid discussion on the subject.[57–60]

However, things for the Kdp ATPase lie slightly differently. The KdpFABC complex is an emergency potassium uptake system, the regulation is not achieved by inactivation but by degradation of the enzyme. Similar to the Ca²⁺-ATPase and the Na⁺,K⁺-ATPase the transmembrane helices TM4 and

TM5 are integrated into the P-domain as part of the Rossmann-fold, but the transmembrane helices TM8 – TM10 were lost in course of evolution. As described before, ion translocation is assumed within the KcsA channel-like subunit KdpA. This means that Kdp is a one-way system only, as no counter-transport of ions is possible. Still, two charged residues (D583 and K586) within TM5 are highly conserved amongst the KdpB proteins of different origin, located precisely at the same place as N768 and E771 of the Ca^{2+}-ATPase (SERCA1a). Mutation of theses residues resulted in an enzyme that was still able to hydrolyse ATP, but it did no longer support ion transport.[61]

The results of this work here show that neither the simplified Post-Albers scheme, nor the E1E2 model according to DeMeis and Vianna hold for a plausible explanation of the reaction cycle of the KdpFABC complex.

Chapter

3

Nuclear Magnetic Resonance

3.1 A Short History

The phenomenon of nuclear magnetic resonance was first discovered in 1946 independently by the two groups around F. Bloch[62] and E. M. Purcell.[63] They were awarded the Nobel Prize in Physics in 1952. Since its discovery NMR spectroscopy has undergone a significant development. An important improvement over the continuous wave spectrometers was invented by Ernst and Anderson in 1966. They used high-power radio frequency pulses, irradiating the whole spectral bandwidth in once and applied a fourier transformation (FT) on the data obtained. The new method signified a huge gain in time. Additionally, the signal of the free induction decay (FID) could be summed up in the computer. This improved the signal to noise ratio significantly. In the beginning of the seventies, Jean Jeener described the first two dimensional correlation spectrum for protons. Unfortunately, his results were only presented as a conference communication, and thus his contribution to the progress of NMR spectroscopy was acknowledged much later when he got the Russel Varian Prize at the EENC

in 2002 in Prague. Again it was Richard Ernst who took over; not without
consulting Jeener first, he combined the 2D technique with FT spectroscopy
and created the basis for modern NMR spectroscopy.[64] For his continuous
contribution to the development of NMR spectroscopy Ernst was awarded
the Nobel Prize in Chemistry in 1991. In the years following Ernst's dis-
covery the development did not stop. The focus was now drawn towards
biomolecules. Here, the group around Kurt Wüthrich set landmarks in the
way to protein structures by NMR[65] which was honoured with the Nobel
Prize in Chemistry in 2002. Again in 2003 the Nobel Prize in Medicine was
awarded to two men who worked on the field of magnetic resonance imag-
ing. Paul C. Lauterbur was the first to present a two-dimensional picture
of a water filled vial inserted into a magnetic field. Peter Mansfield further
developed the technique of introducing gradients in the magnetic field and
elaborated a mathematical procedure to analyse the signals more efficiently
and thus to speed up the imaging process. Today magnetic resonance imag-
ing (MRI) is widely used in medicine as a non invasive and patient friendly
tool for scanning various parts of the body in order to trace disruptions of
the tissue.

3.2 Protein Structure Determination by NMR

3.2.1 Prerequisites – The Magnetic Phenomenon

Matter is constructed of atoms. Atoms themselves are constructed of nu-
cleons (neutrons and protons) and electrons. All atomic particles possess a
spin angular momentum which is intrinsic to them. The sum (or difference)
of the single spin angular momentums of all particles within the atom yields
the net spin. In many cases the overall sum is zero. Nevertheless, in some
cases the net spin of a certain isotope does not cancel out. These isotopes are
sensitive to an external magnetic field. Like the needle of a compass they get
oriented, but, according to their azimuthal quantum number M_S they can

only occupy definite states. An isotope with spin $I = 1/2$ is degenerated twice and its azimuthal quantum number is $M_S = \{-1/2, +1/2\}$. Table 3.1 lists a selection of nuclei that are frequently used for NMR-experiments.

isotope	spin	nat. abundance [%]	magnetogyric ratio γ [10^7 rad/Ts]	rel. sensitivity
^{1}H	1/2	99.985	26.752	1.000
^{2}D	1	0.015	4.107	0.010
^{13}C	1/2	1.10	6.728	0.016
^{15}N	1/2	0.366	-2.712	0.001
^{19}F	1/2	100	25.181	0.833
^{31}P	1/2	100	10.841	0.066

TABLE 3.1: Frequently used isotopes in NMR spectroscopy and their properties.[66]

3.2.2 Assignment Strategies

Once a nucleus is introduced into a magnetic field B_0 and its equilibrium is disturbed by a temporary magnetic field B_1 perpendicular to B_0, it immediately starts to restore equilibrium conditions. The longitudinal or spin-lattice relaxation T_1 brings the spin ensemble back to the Boltzmann equilibrium. It is dependent on temperature and viscosity of the sample and for organic molecules its value is usually in the range of milliseconds to seconds. The transversal or spin-spin relaxation T_2 describes the gradual loss of synchronisation between the precessing nuclear spins. It covers a broader range, reaching from seconds for small molecules to milliseconds for large molecules. It is extremely sensitive to changes in the magnetic field. Both, T_1 and T_2, are extremely valuable in studying backbone dynamics of proteins, as the relaxation times of the individual nuclei are highly dependent on their freedom of movement.[67]
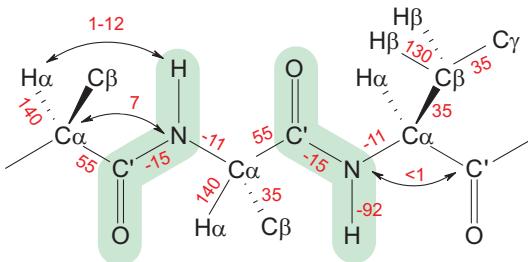
FIGURE 3.1: **Typical values of scalar coupling constants in proteins.**[68,69] The actual size of an individual coupling constant is dependent on the conformation of the molecule. Values are given in Hz. Planar parts are lined green.

Relaxation does not necessarily lead to a loss of magnetisation. Coupled spins have the possibility to transfer coherence to other nuclei either via a scalar coupling, mediated through bonds or via dipolar coupling through space. Figure 3.1 illustrates the typical values for intramolecular J-coupling constants in an uniformly ^{15}N, ^{13}C-labelled protein.[68,69] As the transfer rate of coherence is dependent on the coupling constant it is possible to correlate selected nuclei. The values given in Figure 3.1 are absolute values. To calculate the transfer delays used in practice, relaxation losses have to be taken into account. Due to these relaxation losses the point of maximal transfer is reached slightly earlier. As a rule of thumb, approximately 10 % are substracted from the calculated optimum. The H^N–^{15}N J coupling constant is -92 Hz, thus a typical coherence transfer delay in an ^{1}H,^{15}N-HSQC, which has its calculated maximum at $1/(4J)$ would be 2.7 ms. Indeed, in practice the delay chosen would rather be 2.5 ms. Transfer steps, exploiting the scalar couplings of heteronuclei, are the basis of all double and triple resonance experiments that are nowadays used for protein structure determination.[69]

More details on this topic can be found in the textbooks of J. Evans[70] and M. Levitt.[71]

3.2.3 NOESY-spectra – Key to 3 Dimensional Structures

The nuclear Overhauser effect (NOE) describes the transfer of magnetisation between non J-coupled nuclei. The interaction between two or more nuclear magnetic dipoles through space stimulate cross-relaxation and thus saturation of the transitions between the coupled spins. The intensity of the NOE signal is determined by the internuclear distance and is proportional to the factor $1/r_{1,2}{}^6$. As the NOE build-up is time dependent the magnetisation is "handed on" from one nucleus to another such that after a while the magnetisation is equally distributed over the whole molecule. This process is called spin diffusion. In NOESY experiments, serving for the evaluation of distance restraints, mixing times have to be carefully chosen to ensure a sufficient time for the NOE build-up, but at the same time short enough to avoid spin diffusion. NOESY experiments provide valuable distance information for the identification of three dimensional structures. One severe problem is the rather poor spectral distribution of ^{1}H resonances. In combination with the selection of specific nuclei via scalar couplings, thus an HSQC-type spectrum, the problem of overlap of NOE cross peaks can be reduced (Figure 3.2).

One possibility to overcome the problem of overlap is the application of heteronuclear edited 4D spectra, though sampling in a fourth dimension is very time consuming and may pose a problem when applied to fast relaxing macromolecules. Commonly recorded 4D spectra are the HNNH-, HCCH- and HCNH-NOESY experiments. An efficient approach to avoid the drawbacks of 4D experiments and still keep their advantages, namely the larger dispersion of the chemical shifts of ^{13}C and ^{15}N nuclei, is to leave out the incrementation time in the proton dimension of the first HSQC step. This results in the much simpler, but almost as informative 3D experiments termed NNH-,[72] CCH-, NCH- and CNH-NOESY.[73]
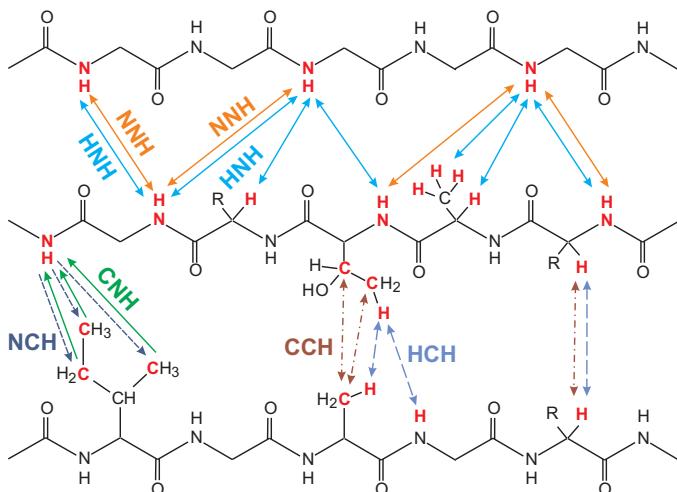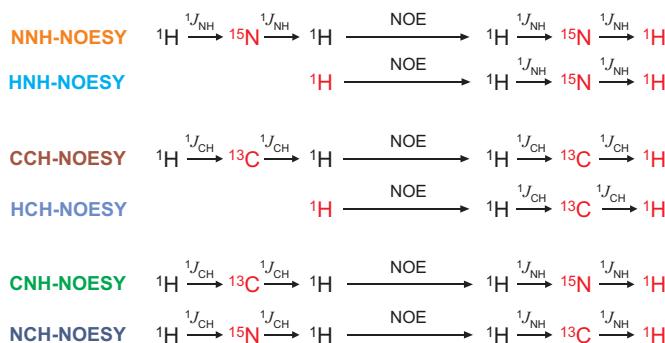
FIGURE 3.2: **Coherence transfer via the nuclear Overhauser effect.** Depicted are the pathways of magnetisation transfer in a selection of commonly used NOESY experiments. Nuclei where time was incremented are shown in red on the upper panel. Examples for the possible structural information yielded therefrom are depicted on the lower panel for an excerpt from an antiparallel β-sheet.

Chapter
4

Structure Determination of the Nucleotide Binding Domain of KdpB

4.1 Protein Expression and Labelling

At the beginning of this project, little was known about the structural properties of P-type ATPases. Some typical sequence motives had been identified[22] and cryo-EM pictures of the Ca^{2+}-ATPase at 8 Å resolution[49] were available, supporting the idea of a modular design of P-type ATPases. Evolutionary studies of Axelsen & Palmgren[21] grouped the Kdp system together with the heavy metal transporting ATPases in type I ATPases. When Toyoshima and coworkers[23] eventually solved the structure of the Ca^{2+}-ATPase by X-ray crystallography at 2.6 Å resolution in the $E1 \cdot Ca^{2+}$ state this signified a milestone in the field of ATPase research.

Now a computer based modelling of the KdpB subunit was possible. Figure 4.1 shows the model of KdpB based on the structure of the Ca^{2+}-ATPase in the $E1 \cdot Ca^{2+}$ state. The modeling of KdpB was performed using the programs "WHAT IF"[74] and "O"[75] by Dr. S. Engelbrecht and Dr. M. Gaßel
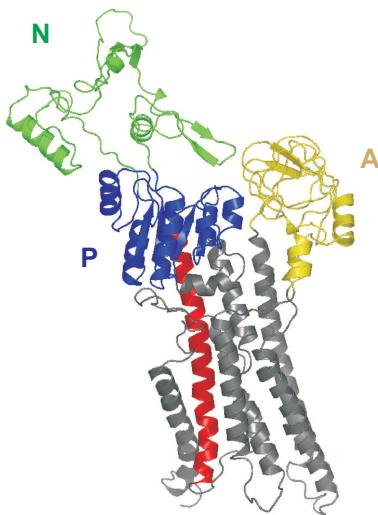
FIGURE 4.1: **Model of KdpB according to the Ca²⁺-ATPase.** Applied were the structural data of the E1·Ca²⁺ state of 1EUL.[23] The domains are coloured yellow (A-domain), blue (P-domain) and green (N-domain). Transmembrane helix 5 was coloured red. (pdb data file obtained from Drs. Engelbrecht, Gaßel and Bramkamp, personal communication)

of the University of Osnabrück (personal communication).

The phosphorylation domain (P-domain) of KdpB shows a good structural alignment with the structure of the Ca²⁺-ATPase. It has approximately the same size and the motive of the phosphorylation site is highly conserved. In contrast, the nucleotide binding domain shows no structural alignment at all. The N-domain of KdpB is approximately 100 amino acid residues shorter than the N-domain of the Ca²⁺-ATPase and the nucleotide binding motive shows a slightly different pattern (KGSVD instead of KGAPE, see also section 2.2).

Cloning of the complete H4H5 loop (32.6 kDa), comprising both the phos-

phorylation and the nucleotide binding domain resulted in a rapidly de-
graded protein; the degradation was apparently taking place within the
cells. In consequence a smaller and more stable construct was required for
NMR analysis. The modular design presented the possibility of cutting out
the nucleotide binding domain. Subcloning of the corresponding residues
N^{316}–G^{451} into a pET16b vector gave a deca-histinyl fusion protein, which
was much more stable. The computed M_w for the protein including the deca-
histinyl tag was 17134.23 Da (136+20 residues). The result of expression and
purification of KdpBN is shown in a SDS polyacrylamide gelelectrophoresis
run (SDS-PAGE) in Figure 4.2. A detailed description of cloning and purifi-
cation steps can be found in refs. [25,34].
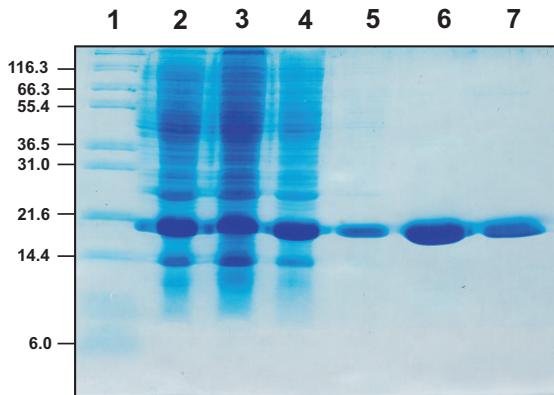

FIGURE 4.2: **Coomassie blue G250 stained SDS-PAGE of purified of KdpBN.** (1)
protein marker, (2) induced cells, (3) cytoplasmic extract, (4) flow through, (5) last
washing step, (6) elution 1, (7) elution 2. Elution 1 was concentrated and used for
NMR analysis. Elution 2 and the solution of the last washing step were unified and
purified again. (Marc Bramkamp, personal communication)

To ensure that KdpBN was still functional, FITC binding was performed on the KdpFABC complex and the constructs of the H4H5 loop and the N-domain of KdpB. FITC binds covalently to the protein, forming a condensation adduct to lysines involved in nucleotide binding. As shown in Figure 4.3 all constructs bind FITC well, whereas denatured KdpFABC complex (SDS) does not, ruling out the possibility of unspecific binding. In case of the complete H4H5 loop and the KdpFABC complex a protection of the binding site can be observed when the proteins were preincubated with ATP, ADP or AMP even after incubating them afterwards for half an hour with a FITC solution. Even when applying only short incubation times, this protectional effect is no longer given for KdpBN, suggesting a rapid exchange of the adenosine nucleotides.[25]
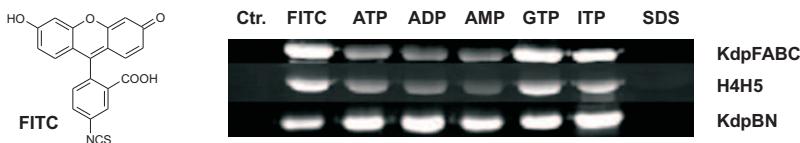
FIGURE 4.3: **FITC binding to the KdpFABC complex, the H4H5 loop and KdpBN.** Purified KdpFABC complex, H4H5 loop and KdpBN were preincubated with 5 mM of the different nucleotides as indicated above, before being treated with 15 mM FITC for 30 min at 37°C. The FITC modification was stopped by denaturation and the mixture was applied to SDS-PAGE. The gels were visualised under UV-light (366 nm). SDS denatured protein was used as control for specific binding (right lane). A fluorescent control with untreated protein (Ctr.) was added to show that the protein does not exhibit autofluorescence. (Marc Bramkamp, personal communication)

Furthermore it could be shown that KdpBN still remains functional when it is concentrated in potassium phosphate buffer above 1 mM at pH 6. First ^{1}H,^{15}N-HSQC spectra of uniformly ^{15}N-labelled protein showed a good distribution of the signals, confirming that the nucleotide binding domain is well structured. Best stability was obtained when using a 50 mM potassium phosphate buffer with 100 mM NaCl.
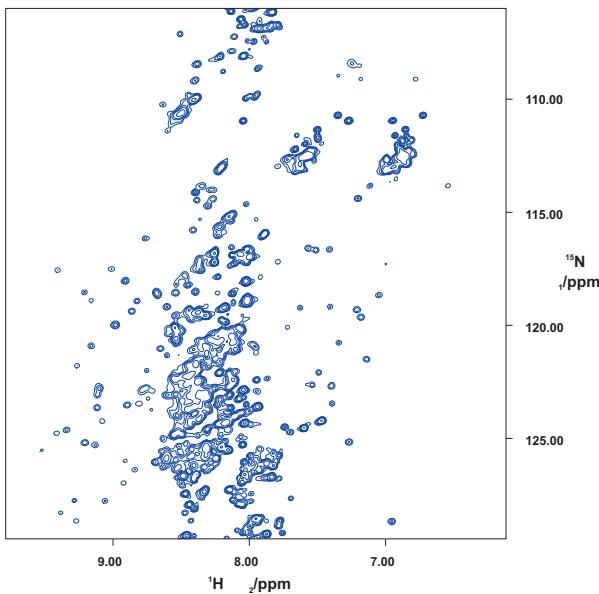
FIGURE 4.4: **^{1}H,^{15}N-HSQC from double labelled KdpBN from bacteria grown in *Silantes OD4 CN* medium.** The majority of the protein is not folded, though in solution.

When preparing the uniformly ^{13}C,^{15}N-labelled samples it seemed to be conclusive to change from minimal medium to rich medium, as the expression rates in rich medium are significantly higher than in minimal medium. For the labelling experiments a commercially available rich medium *Silantes OD4 CN* from Silantes (München, Germany) which contains only ^{13}C and ^{15}N labelled substances was chosen, as previous experiments with this medium had shown a much better overproduction of protein. OD4 refers to the maximal optical density up to which bacterial growth in the medium can be put forth. The yield of protein was indeed higher than it had previously been in minimal medium, but the protein proved not to be folded, though

soluble. The resulting ^{1}H,^{15}N-HSQC spectrum is shown in Figure 4.4. Only a minor part of the protein assumes the correct fold (small, well distributed signals), whereas the majority of the protein is unfolded and thus the amide proton signals show the typical chemical shifts for a random coil protein between 8.0–8.5 ppm.

After three days the protein started to precipitate. To renature the protein, the double labelled protein suspension was diluted to approximately 1 mg/mL (0.06 mM) in 50 mM potassium phosphate buffer at pH 6 containing 1 % (27 mM) octyl glycoside, 5 mM EDTA and 5 mM β-mercaptoethanol.[76] Octyl glycoside (Figure 4.5) is known to be a very gentle, "protein friendly" detergent as is does not bear any charges. After stirring the suspension at room temperature for 3 hours the protein was completely resolubilised. It was stirred for another hour at 4°C to complete solubilisation, then it was dialysed against buffer over polyurethan beads and eventually reconcentrated to 1.0 mM. Unfortunately, all rescue attempts were in vain. The protein did not regain its native fold.
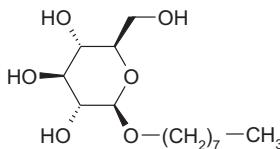
FIGURE 4.5: **Octyl glycoside.** A frequently used detergent for protein renaturation.

To investigate on the reasons why the protein was not folded when expressed in *Silantes OD4 CN* medium a row of experiments was set up. Firstly, the expression of KdpBN was repeated in unlabelled *Silantes OD4* medium. Secondly, the cells were grown in Luria Bertani (LB) rich medium[77] to have another rich medium for comparison. Thirdly, KdpBN was expressed in minimal medium (K115) with 1 % glucose. The overexpressed protein from the three expression systems were equally treated in the following process
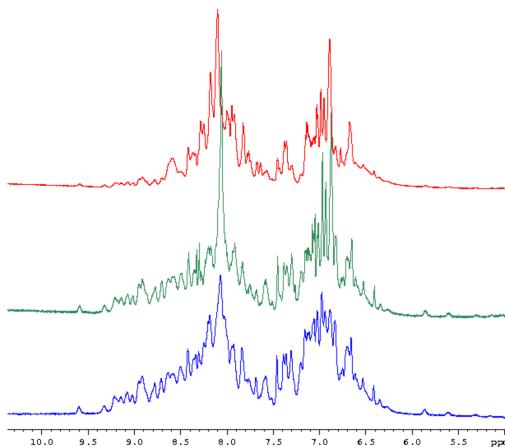
FIGURE 4.6: **1D proton spectra of unlabelled KdpBN from bacteria grown in different media.** (red) *Silantes OD4* medium, (green) LB rich medium, (blue) minimal medium with 0.2 % glucose. The protein obtained from the *Silantes OD4* medium is unfolded, whereas KdpBN expressed in LB rich medium and minimal medium is folded.

of purification and concentration.

1D proton spectra were recorded to compare the proteins. Interestingly, the expression of unfolded but soluble protein in *Silantes OD4* medium was reproducible. Figure 4.6 shows the amide proton region of the 1D proton spectra. The dispersion of the signals is clearly worse in the case of the protein expressed in *Silantes OD4* medium. The signals at the outer ends of the scale, e.g. at 5.6, 5.9 and 9.6 ppm are almost invisible, which is a clear indication that no tertiary structure is formed.

For future protein expression minimal medium was used throughout all experiments. Cells were either grown with 1.0 % glucose in case of ^{15}N-labelleling with $(^{15}NH_4^+)_2SO_4^{2-}$ or 0.2 % [U-$^{13}C_6$]-glucose in case of double labelled protein.

4.2 Resonance Assignment and Secondary Structure Prediction

The sequential assignment of the backbone resonances of KdpBN was achieved using the automatic assignment program PASTA.[78] Therefore approximately 180 peaks were picked from a well resolved ^{1}H,^{15}N-HSQC spectrum representing the basic input. Following the carbonyl carbon resonances in $(i-1)$ position were assigned using an HNCO experiment. Then a combination of a CBCA(CO)NH, HNCA, HNCACB, HBHA(CO)NH, HACACO and HN(CA)HA were used to assign the C^α, C^β, H^α and H^β chemical shifts of residue i and the corresponding residue $(i-1)$. The resonances were entered into the PASTA spreadsheet, then the program was launched. PASTA searches iteratively for the be residues that are in best agreement of the residues in i and $(i+1)$ position. Depending on the accuracy of the data a maximal deviation of the chemical shifts can be fixed and a total energy threshold gives an idea of the exactitude of the assignment.

Finally, side chain assignment of the aliphatic residues was done using an (H)CCH-COSY experiment. Glutamine and asparagine side chain NH$_2$ resonances were assigned using the CBCA(CO)NH experiment. The ^{1}H chemical shifts of the aromatic residues could partially be assigned via the distance information of the CCH-NOESY experiment. Therefore the contacts of C^β are closely examined. In case of phenylalanine medium strong NOE cross peaks can be observed from H^β to H^δ. Slightly weaker are the contacts from H^β to the aromatic H^ϵ. H^ζ can only be found in rare cases as it is already more than 4 Å away and it usually assumes an unfavourable angle towards H^β. However, an unambiguous assignment of all aromatic residues could not be achieved. An aromatic ^{1}H,^{13}C-HSQC experiment revealed that the ^{1}H frequencies of the three phenylalanines in KdpBN are all resonate in the area of 7.12–7.38 ppm, which explains the difficulties in deriving the data. Figure 4.7 shows the fully assigned ^{1}H,^{15}N-HSQC spectrum of KdpBN recorded at 750 MHz.
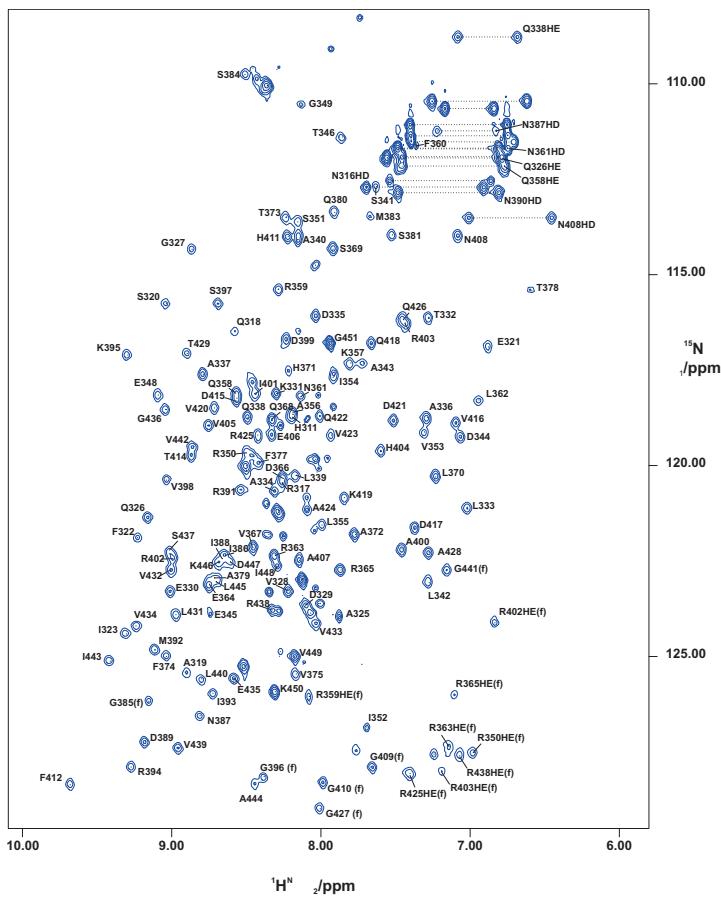
FIGURE 4.7: **^{1}H,^{15}N-HSQC of KdpBN.** Complete assignment of the backbone reso-
nances and almost complete side chain assignment could be achieved.[79] (f) as a
suffix to peak labels means folded, once in case of glycine (G), twice in case of argi-
nine (R) H$^\epsilon$. Cross peaks connected by dotted lines correspond to side chain NH$_2$
groups of asparagine (N) and glutamine (Q) residues. Unassigned peaks derive from
the N-terminal His-taq or unassigned side chain resonances.
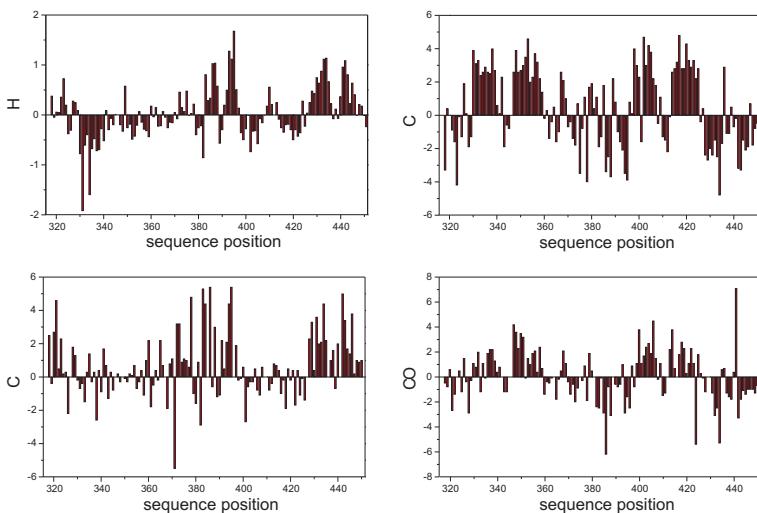
FIGURE 4.8: **Secondary chemical shifts for KdpBN.** Shown are the deviations from the random coil values for H^α, C^α, C^β and C'.

The assignment of the backbone and C^β resonances makes a preliminary estimation of the secondary structure possible. The chemical shifts of H^α, C^α, C' and C^β are compared to the chemical shifts of the corresponding amino acid residues in an unstructured protein chain. The deviation from the random coil values (the secondary chemical shifts) of KdpBN for the atoms in question is shown in Figure 4.8. Still, an interpretation based only on the relative deviations is critical.

It has been known since the early days of protein spectroscopy that secondary structure and chemical shifts are closely linked.[80] It could be shown that C^α and C' resonate downfield when located in an α-helix and upfield when in β-strands; C^β and H^α behave contrarily. The random coil chemical shift values have been described in detail by Wishart & Sykes.[81,82] They suggested an easy approach to define the secondary structure with
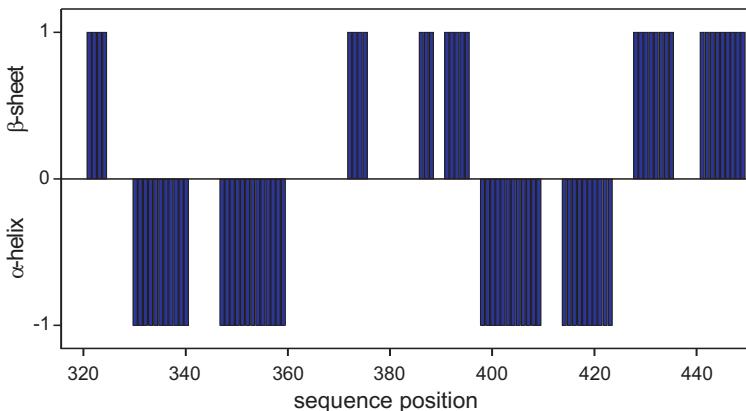
FIGURE 4.9: **CSI consensus plot of KdpBN.** The evaluation of the secondary chemical shifts with the CSI-program shows the protein to have six β-strands and four α-helices.

high accuracy by defining the chemical shift index, CSI. There the values of the secondary chemical shift are classified in either $+1$, when the chemical shift, less a certain tolerance, is greater than the random coil value of the corresponding atom, or -1, when the chemical shift observed is less. The consensus CSI summarises the results of the independent approaches from the different nuclei. α-helices assume per definition the value -1, whereas β-strands are described by $+1$. In consequence the values for C^{β} and H^{α} are multiplied by -1 before being merged with the CSI of C^{α} and C'. The existence of an α-helix is suggested, when a dense grouping of at least four -1's is not interrupted by a $+1$. For the formation of a β-strand three consecutive $+1$'s are sufficient. The resulting CSI consensus plot for KdpBN is shown in Figure 4.9.

In order to get reproducible results from the CSI analysis, the chemical shifts have to be carefully referenced, as they are sensitive to pH and buffer conditions. In the case of KdpBN, a few milligrams of DSS (4,4-dimethyl-

4-silapentane-1-sulfate4,4-dimethyl-4-silapentane-1-sulfate) were added to a sample that was no longer required. The ^{1}H chemical shift was directly referenced to DSS and the chemical shifts of ^{13}C and ^{15}N were indirectly referenced.[83]

Another approach to define the secondary structure of a protein is the evaluation of the homonuclear $^3J(H^N,H^\alpha)$-coupling constants.[84] The size of the coupling constants is dependent on the dihedral angle ϕ. It can be described by the following Karplus equation:[71]

$$^3J_{\text{Ha,NH}}(\phi) = 6.4\cos^2(\phi - 60°) - 1.4\cos(\phi - 60°) + 1.9$$

Taking into consideration that only certain combinations of ϕ and ψ angles are allowed in secondary structured elements, it follows that only certain coupling constants can be true for a certain conformation. These conclusions are condensed in the Ramachandran plot. The typical ϕ angle of a perfect α-helix is $-57°$, this results in a $^3J(H^N,H^\alpha)$ coupling constant of 3.9 Hz. The values for an antiparallel β-sheet are $\phi = -139°$, and thus $^3J(H^N,H^\alpha) = 8.9$ Hz.[70]

The $^3J(H^N,H^\alpha)$ coupling constants cannot be measured directly, but they are determined indirectly via the ratio of the cross peaks to the diagonal signal in the HNHA experiment, following the equation given by Vuister & Bax.[84]

$$S_{cross}/S_{diag} = -tan^2(2\pi J_{HH}\zeta)$$

Where S_{cross}/S_{diag} is the relation of the intensities between the cross peak and diagonal signals and 2ζ is the time span during which the homonuclear H^N–$H^\alpha J$ coupling evolves. In standard Bruker pulseprograms ζ is named d3. To get an optimal transfer, it was set to 14 ms. The coupling constant can be calculated via the following equation.
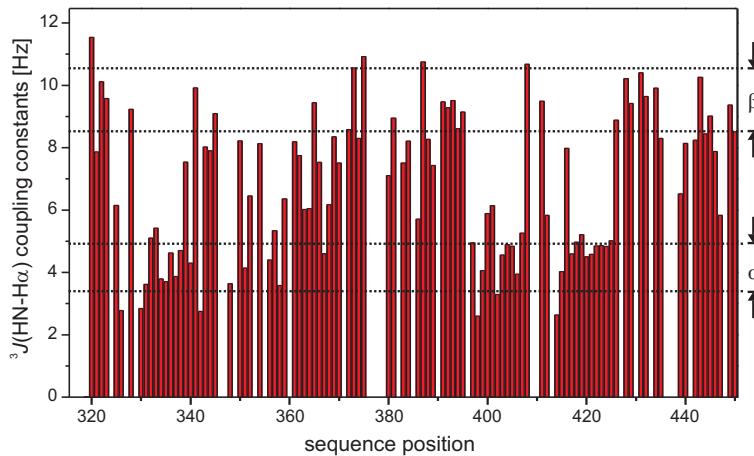
FIGURE 4.10: $^3J(\mathbf{H^N,H^\alpha})$**-coupling constants.** The regions which have been assigned to be α-helical or β-strand dominated are enclosed by dashed lines.

$$^3J(\mathrm{H^N,H^\alpha}) = \frac{arctan\sqrt{-(S_{cross}/S_{diag})}}{2\pi\,\zeta\,e}$$

As the $\mathrm{H^N}$–$\mathrm{H^\alpha}$ cross peaks result from a faster relaxing anti-phase term and the $\mathrm{H^N}$–$\mathrm{H^N}$ diagonal peaks from the slower relaxing in-phase term, the relative propensities are corrected by the factor e. Usually, e is in the order of magnitude of ≥ 0.9. In case of KdpBN, it was adjusted to 0.8 to get reasonable results for the coupling constants, which are now in the range of 0.0–11.6 Hz.

To compare the structure information derived from the $^3J(\mathrm{H^N,H^\alpha})$-coupling constants with the CSI, the coupling constants were transformed into a similar index. To clarify the representation, the values were multiplied by two. In consequence, $+2$ describes an coupling constant of $(8.9 \pm 0.5)\mathrm{Hz}$ and thus a β-strand; -2 stands for a coupling constant of $(4.0 \pm 0.6)\mathrm{Hz}$ and
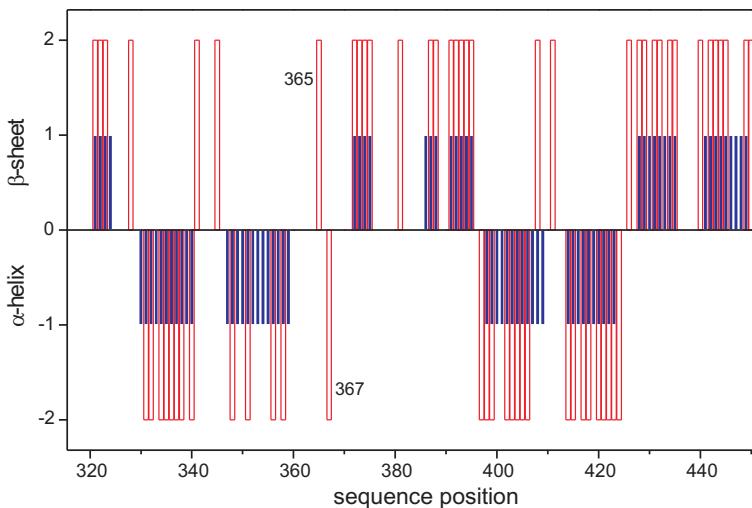
FIGURE 4.11: **Secondary structure of KdpBN.** The consensus plot of the chemical shift index (CSI, solid blue) and the secondary structure prediction via the homonuclear $^3J(H^N,H^\alpha)$ coupling constants (hollow red) are in good agreement. Both predict six β-strands and four α-helices to shape the protein. The values for the $^3J(H^N,H^\alpha)$ coupling constants were set to 2 or -2 for the sake of clarity. Residues 365 and 367 show coupling constants typical for a β-strand or α-helix, respectively. This indicates that a short 3_{10}-helix might be present at this position.

thus an α-helix. The result is shown in Figure 4.11. Both independent methods show a good agreement in the predicted secondary structure. The shape of the protein is characterised by six β-strands and four α-helices. Only α-helix 2 seems to be less well defined. The residues 365 and 367 show an unusual pattern in the evaluation of the $^3J(H^N,H^\alpha)$-coupling constants. The sudden switch from β-strand-like to α-helical-like angles within three residues is typical for a short 3_{10}-helix.

4.3 The Tertiary Structure of KdpBN

A first rough glimpse can be thrown on the tertiary structure when evaluating the 3D-^{15}N-HSQC-NOESY-^{15}N-HSQC experiment, short NNH-NOESY. As the dispersion in the amide region is good, an unambiguous assignment of most H^N–H^N contacts is already possible. In case of ambiguities, e.g. such as very weak cross peaks, the results of the NNH-NOESY can be compared to the 3D-^{15}N-HSQC-NOESY experiment, short HNH-NOESY, which has basically the same information in the amide region plus the correlations to the aliphatic protons. First of all, only cross peaks connecting the β-strands are sought, as they provide the information about the topology of the protein. The H^N–H^α contacts across the β-strands serve as a confirmation. For an antiparallel β-sheet, as is the case for KdpBN, an inter-strand H^N–H^N NOE cross peak from (i) to (i') necessarily has to be accompanied by a strong H^N–H^α cross peak from (i) to ($i' + 1$) (see Figure 4.12). Typical for antiparallel β-sheets are extremely strong H^α–H^α cross peaks. As they are directly facing each other their distance is below 2.2 Å in an ideal antiparallel β-sheet.
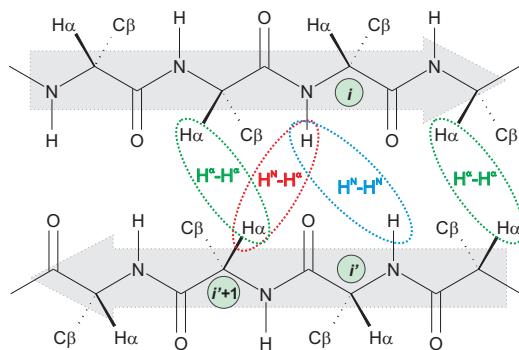
FIGURE 4.12: **Connectivities in an antiparallel β-sheet.**
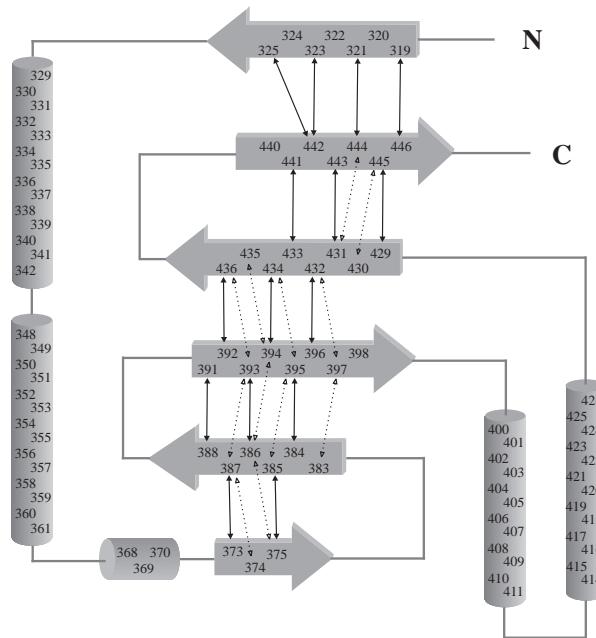
FIGURE 4.13: **The topology of KdpBN.** Shown are the H^N–H^N connectivities, derived from an NNH-NOESY experiment (filled lines) and the H^N–H^α contacts from the HNH-NOESY experiment (dotted lines).

Figure 4.13 illustrates the combination of the secondary structure information obtained from the CSI with the first, unscaled NOE restraints from the NNH- and HNH-NOESY spectrum. These data also serve as the first input for a structure calculation.

Following the assignment of NOE connectivities were accomplished using a combination of complementary HSQC based 3D-NOESY experiments as described in section 3.2.[73] As an optimal NOE mixing time a value of $\tau_{mix} = 80$ ms was determined; a time span where the relaxation of the

17.1 kDa protein was still in tenable limits and spin diffusion did not pose a problem yet.

Best information could be yielded from the CNH-NOESY spectrum for tightening the constraints and finally the HCH-NOESY provided data concerning the side chain arrangements. A CCH-NOESY spectrum provided information where the HCH-NOESY failed due to overlap with the water signal. The cross peaks were divided into four classes, strong, medium, weak and very weak, which resulted in restraints on upper distances of 2.7, 3.2, 4.0 and 5.0 Å, respectively. For the NOESY spectra recorded at the 900 MHz spectrometer, the scaling of the distance restraints was adapted to the increased sensitivity by setting the upper distance restraints to 2.7, 3.2, 4.2 and 5.5 Å, respectively.
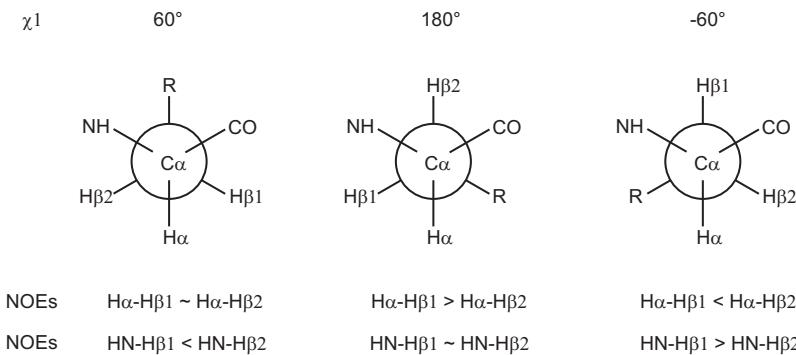
FIGURE 4.14: **Scheme for obtaining stereospecific assignments of H^β.**[71] The distance information can be taken from an HNH-NOESY spectrum by comparing the relative intensities of the cross peaks in question.

Parallel to the evaluation of the NOESY spectra stereospecific assignment of the H^β atoms was carried out. The assignment of most β-methylene protons could be achieved by evaluating the NOE cross peaks from H^N to $H^{\beta 1}$ and $H^{\beta 2}$ and comparing them to the cross peaks from H^α to $H^{\beta 1}$ and $H^{\beta 2}$. As

only three allowed rotamers exist, the evaluation is usually straightforward (see Figure 4.14).

In total 2232 NOE distance restraints were evaluated, which makes approximately 18 restraints per structured residue. 433 of these NOE restraints were intra-residual contacts, 674 were derived from sequential inter-residue contacts. 30 % of all restraints (676) were long range NOEs and 20 % (449) medium range. The high number of long and medium range restraints forms the basis for the well defined structure of KdpBN. The calculated structure has an RMS deviation of 0.17 Å for superimposition over ordered backbone residues (see Table 4.1) and an RMSD of 0.41 Å for superimposition of all residues, including unstructured regions and side chains. No dihedral ϕ and ψ angles of the final set of 20 structures occur in the disallowed region of the Ramachandran plot (Figure 4.15).

The N-domain forms a curved, six-stranded, antiparallel β-sheet flanked by two pairs of α-helices, one on the concave side of the sheet and tightly packed to form the hydrophobic core of the domain, the other more solvent exposed on the convex side (Figure 4.16). In other P-type ATPases the loop between helix α2 and the β2 strand contains an additional α-helix. In KdpBN this is replaced by a short 3_{10}-helix at the C-terminal end of the loop. Beside the C-terminus, only the loop between strand β2 and β3 shows considerable flexibility.

A. Structural statistics

RMSD from distance restraints (Å)[I]	SA	$<SA>_r$
all (2232)	0.025 ± 0.0004	0.025
inter-residue (433)	0.019 ± 0.0014	0.018
inter-residue sequential (674)	0.025 ± 0.0007	0.027
medium range (449)	0.026 ± 0.0004	0.026
long range (676)	0.028 ± 0.0006	0.028
RMSD from dihedral restraints (256)	0.376 ± 0.017	0.358
RMSD from *J*-coupling restraints (Hz)(67)	0.572 ± 0.012	0.572
H-bond restraint violations (Å/deg)[II] (66)	$2.04 \pm 0.20/17.7 \pm 7.4$	2.02/16.3
Deviations from ideal covalent geometry		
Bonds (Å$\times 10^{-3}$)	3.01 ± 0.003	3.00
Angles (deg)	0.592 ± 0.002	0.594
Impropers (deg)	1.86 ± 0.029	1.88
Structure quality indicators[III]		
Ramachandran map regions (%)	90.8/ 9.2/0.0/0.0	92.5/7.5/0.0/0.0
Bad contacts per 100 residues	8.3 ± 1.1	5.1

B. Atomic R.M.S. differences (Å)[IV]

	SA vs $<SA>$		SA vs $<SA>_r$	
	Backbone	All	Backbone	All
All residues	0.41 ± 0.161	0.78 ± 0.123	0.54 ± 0.171	1.02 ± 0.163
Ordered residues[V]	0.17 ± 0.050	0.61 ± 0.052	0.26 ± 0.061	0.81 ± 0.046
$<SA>$ vs $<SA>_r$[VI]	0.20	0.61		

TABLE 4.1: **Structural statistics and atomic R.M.S. deviations.** Structures are labelled as follows: SA, the set of 20 final simulated annealing structures; $<SA>$, the mean structure calculated by averaging the coordinates of SA structures after fitting over secondary structure elements; $<SA>_r$, the structure obtained by regularising the mean structure under experimental restraints. [I] Numbers in brackets indicate the number of restraints of each type. [II] H-bonds were restrained by treating them as pseudo-covalent bonds (see chapter Methods and Experiments section 6.3). Deviations are expressed as the average distance/average deviation from linearity for restrained H-bonds. [III] Determined using the program PROCHECK.[85] Percentages are for residues in allowed/additionally allowed/generously allowed/disallowed regions of the Ramachandran map. [IV] Based on heavy atoms superimpositions. [V] Defined as residues 318-361 and 366-446. [VI] RMS difference for superimposition over ordered residues.
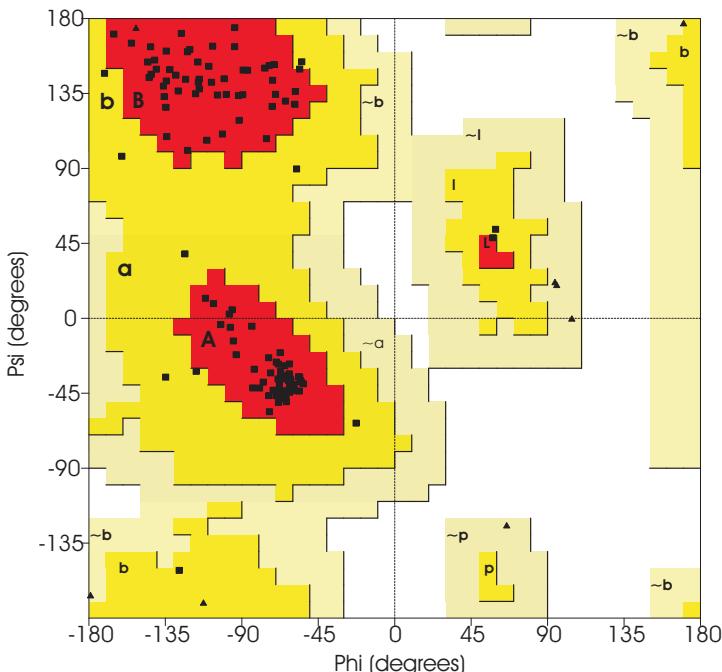
Plot statistics

red	most favoured regions [A, B, L]	111	95.5 %
yellow	additional allowed regions [a, b, l, p]	9	7.5 %
light yellow	generously allowed regions [a~,b~,l~,p~]	0	0.0 %
white	disallowed regions	0	0.0 %
	Nº of non-Gly and non-Pro residues	120	100 %
	Nº of end-residues (excl. Gly and Pro)	1	
	Nº of Gly residues (shown as triangles)	10	
	Nº of Pro residues	5	
	Total Nº of residues	136	

FIGURE 4.15: **Ramachandran plot of the regularised mean structure.** For glycine residues (black triangles) this classification is not representative.[85]
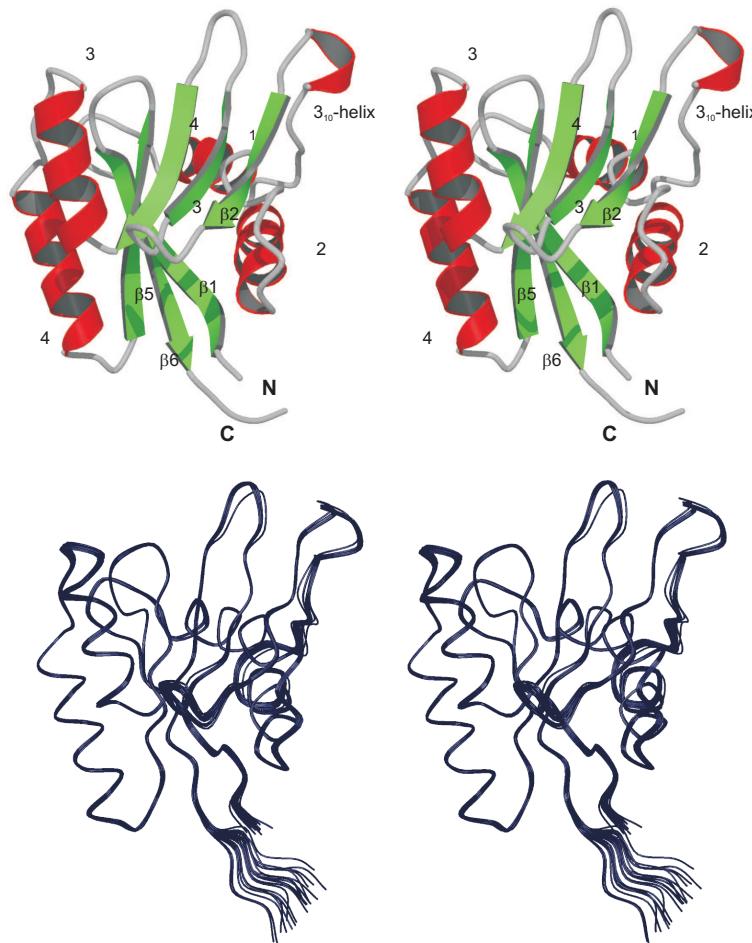
FIGURE 4.16: **Stereo view of the structure of KdpBN.** The upper panel shows the secondary structure elements. Helix $\alpha1$ lies horizontally behind the β-sheet. The lower panel shows the definition of the final set of 18 structures. The structure set is superimposed over ordered residues (RMSD 0.17 Å).

Chapter

5

Nucleotide Binding Studies on KdpBN

To complete the analysis of KdpBN, further investigations on the nucleotide binding properties were carried out. KdpBN represents the isolated nucleotide binding domain of the KdpFABC complex. Previous biochemical assays had shown that it had retained the native fold as it still binds nucleotides (see section 4.1).

Three major question were addressed in this context. First, where is the binding pocket located? Second, what is happening to the protein upon nucleotide binding? And third, how is the nucleotide placed in it?

The answers to these questions provide insights in the mechanism of the whole KdpFABC complex. Additionally, the evolutionary origin of the Kdp system can be enlightened, as for the moment KdpB is grouped together with the heavy metal transporting P-type ATPases (type I), but in the basic sequence motive at the nucleotide binding site it shows a closer relation to the type II–IV P-type ATPases (compare Figure 2.7).
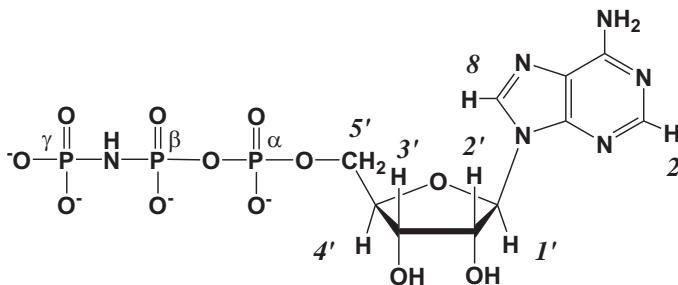
FIGURE 5.1: **Numbering scheme of adenosine 5'-(β,γ-imido)triphosphate.** AMP-PNP is a commonly used, non hydrolysable substitute for ATP.

5.1 TNP-nucleotide Displacement Experiments

To start with, ATP had to be replaced by a stable analogue, as it undergoes fast autohydrolysis when in solution. Even at neutral buffer conditions ATP hydrolyses quickly. Biochemical tests showed that within an hour approximately 10–20 % are degraded when the solution is stored at room temperature. For the studies on KdpBN the buffer conditions were slightly acidic (pH 6) and NMR experiments were run at 300 K. For this reason an even faster autohydrolysis rate can be expected for an ATP–protein mixture. In addition it is possible that the nucleotide binding domain activates the phosphate ester bond and thus stimulates hydrolysis, even though the phosphorylation domain and with it the acceptor of the phosphate group is not present. A common substitute for ATP is adenosine 5'-(β,γ-imido)triphosphate (AMP-PNP), which is often used in NMR spectroscopy.[86–88] In AMP-PNP the phosphate ester bond between the β- and γ-phosphate end group is replaced by an amine group (see Figure 5.1). This increases the stability of the molecule while the structure remains basically the same.

To verify that AMP-PNP has indeed the same binding properties as other nucleotides, 2',3'-O-(2,4,6-trinitrophenyl) nucleotide displacement experi-
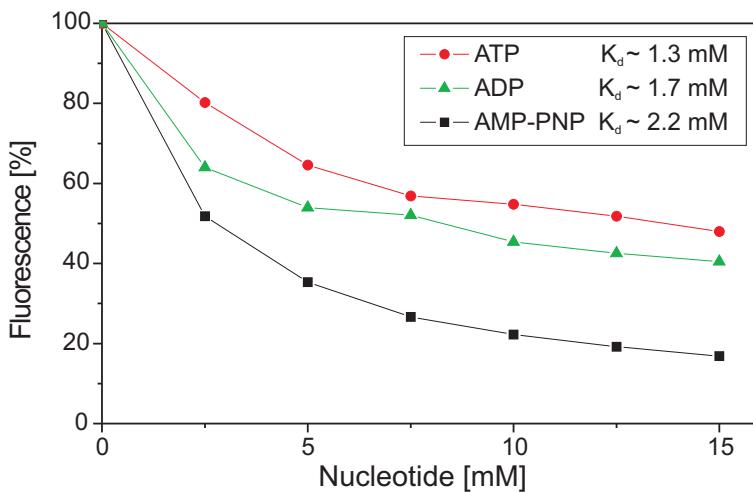
FIGURE 5.2: **TNP-ATP displacement by different nucleotides for KdpBN at pH 6.0.** Measured was the decline of fluorescence of bound TNP-ATP upon addition of nucleotide. Fitting of the curves showed an almost similar displacement constant K_d for all nucleotides.

ments were carried out to determine the K_d values at the given sample conditions. TNP-nucleotides are known to bind specifically to the nucleotide binding sites of P-type ATPases.[89–91] They show a drastic increase in fluorescence at 545 nm upon binding to a protein. A 1 mM solution of protein was preincubated with TNP-ATP. Then a defined amount of ATP, ADP or AMP-PNP was added and the decrease of fluorescence was recorded (data provided by Marc Bramkamp, personal communication). Fitting of the curves resulted in K_d values in the millimolar range for the different nucleotides (see Figure 5.2). ATP and ADP bind with a slightly higher binding constant to the protein. However, the curves showed quite large deviations from the optimal fit line. Thus, the margin of error was estimated to be at least ± 0.3 mM.

5.2 NMR Titration Experiments – Location of the Binding Pocket

In order to identify the location of the nucleotide binding pocket, a 1 mM solution of ^{15}N-labelled KdpBN was titrated with AMP-PNP. As the binding constant of nucleotide to KdpBN is relatively low, the exchange between nucleotide-bound and free state is quick on the NMR time scale. Thus, the resonances of the nuclei affected by nucleotide binding gradually shift their position from the resonance for the free state towards the resonance of the bound state. In most cases this facilitates an unambiguous assignment of the amide proton and nitrogen resonances as they can be traced along their way. In order to avoid changes of the K_d value by dilution effects the changes in volume were kept beyond 10 %. Therefore, the nucleotide was added in small amounts of a highly concentrated solution. As the binding constant is very low, complete saturation is practically unreachable. When relative saturation is reached, the curve changes into a straight line with a slightly positive gradient. This might be due to the fact that at high concentrations of substrates the behaviour of the solution is no longer ideal. The same behaviour had been observed previously with the TNP-nucleotide displacement experiments. For KdpBN, the point where the curve goes over in a straight line, is reached at a ratio of protein to nucleotide of 1 : 7. Thus the titration experiment was stopped at a ratio of 1 : 8. The titration was carried out directly in the NMR tube. To ensure that the nucleotide was readily distributed in the small volume, the protein-nucleotide solution was vortexed for a minute after each titration step before it was reinserted into the magnet. The sample was allowed to equilibrate temperature for 10 min, then the ^{1}H,^{15}N-HSQC was recorded with a resolution of $t_1 \times t_2 = 200 \times 1024$ and 16 scans per increment.

Figure 5.3 shows a superimposition of the ^{1}H,^{15}N-HSQC spectra of native KdpBN and of nucleotide bound KdpBN. A343 and E348 show in addition to a large change in chemical shift an extreme line broadening, indicating
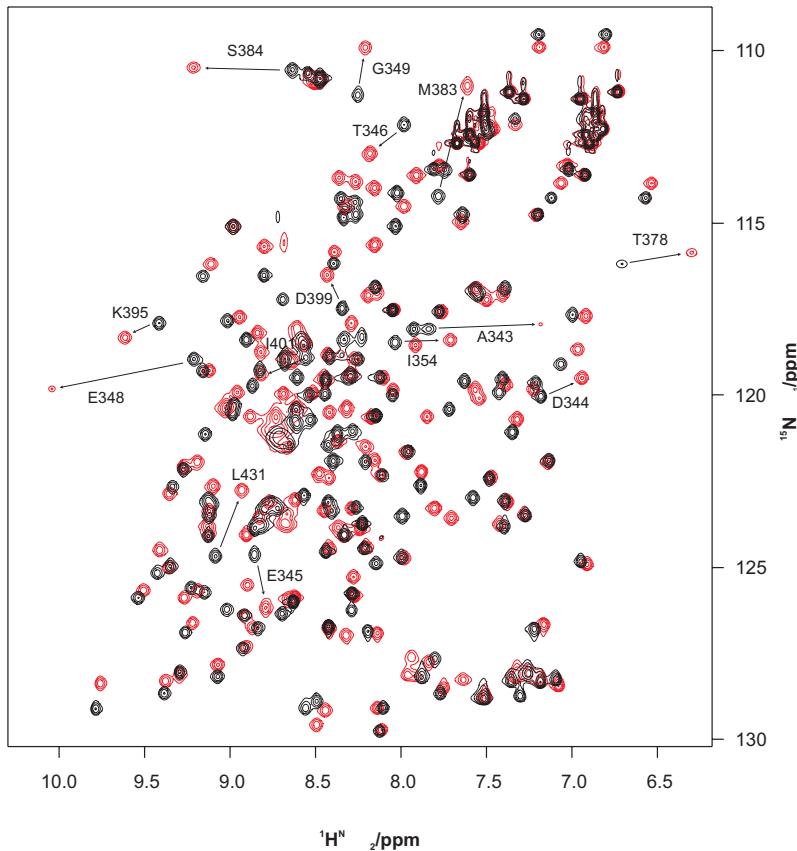
FIGURE 5.3: **^{1}H,^{15}N-HSQC spectra of native KdpBN (black) and AMP-PNP bound KdpBN (red).** The ratio of native to nucleotide bound KdpBN is 1 : 8. The residues that shift most considerably are indicated with arrows. Most of them lie in the loop region between strand $\beta2$ and $\beta3$. K395 suffers a medium sized change in chemical shift upon addition of nucleotide.
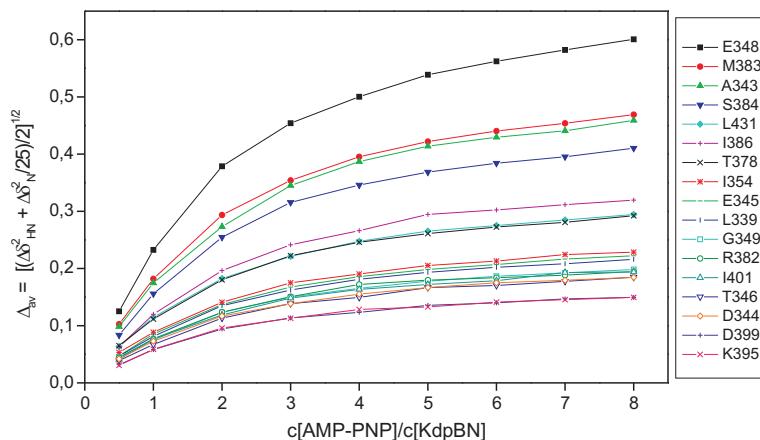
FIGURE 5.4: **The average change in chemical shift (Δ_{av}) upon titration with AMP-PNP.** The plot is only shown for selected residues. Fitting of the curves yielded a calculated K_d value of 1.38 mM.

that they are subject to a medium fast conformational exchange close to the coalescence temperature.

The averaged differences in chemical shifts were calculated according to the empirical equation 5.1.[87,92] This equation takes into account that the spectral width of the amide ^{15}N signals is about 5fold as much as the one of the amide ^{1}H signals.

$$\Delta_{av} = \sqrt{\frac{\Delta\delta_{HN}^2 + \left(\Delta\delta_N/5\right)^2}{2}} \qquad (5.1)$$

A plot of Δ_{av} as a function of the molar ratio of AMP-PNP : protein is shown in Figure 5.4 for selected residues. The shape of the lines is that of a typical saturation curve, only that it does not reach saturation, but rather turns into a straight line with a positive gradient, a point that has already been discussed earlier.

A common model to describe enzyme activity K_M and the rate of catalysis V in biological systems, is the Michaelis-Menten equation (5.2).

$$V = V_{max} \frac{[S]}{[S] + K_M} \tag{5.2}$$

The Michaelis-Menten equation does neglect the concentration of the enzyme. The enzyme is assumed to be present in very low, 'catalytic', concentrations that are far beyond the size of K_M. Thus, the kinetic behaviour is only dependent on the concentration of the substrate. Furthermore, the kinetic model of Michaelis-Menten bases on the assumption that the substrate S gets readily transformed via an intermediate substrate-enzyme complex ES to yield the product P (reaction 5.3).

$$E + S \underset{k_2}{\overset{k_1}{\rightleftharpoons}} [ES] \overset{k_3}{\rightarrow} E + P \tag{5.3}$$
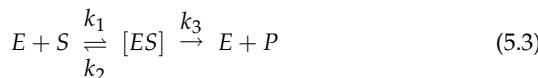
As the NMR titration experiment were recorded at relatively high concentrations of protein, the Michaelis-Menten equation does not longer hold. Here, the concentration of the enzyme cannot be neglected any longer. Additionally, no conversion of substrate takes place as the phosphorylation domain is missing. The reaction that describes the equilibrium of native and nucleotide bound KdpBN is shown in 5.4.

$$[KdpBN] + [AMPPNP] \underset{k_2}{\overset{k_1}{\rightleftharpoons}} [KdpBN \cdot AMPPNP] \tag{5.4}$$
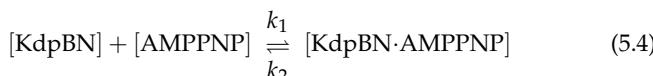
Thus, the dissociation constant K_d for KdpBN is defined by equation 5.5.

$$K_d = \frac{[E][S]}{[ES]} = \frac{[KdpBN][AMPPNP]}{[KdpBN \cdot AMPPNP]} = \frac{k_2}{k_1} \tag{5.5}$$

Further, the concentrations are given by

$$E_0 = [E] + [ES] \quad \text{and} \quad S_0 = [S] + [ES] \,. \tag{5.6}$$

Insertion of equations 5.6 into equation 5.5 and resolving for $[ES]$ results in equation 5.7.

$$[ES]_{1,2} = \frac{K_d + E_0 + S_0 \pm \sqrt{(-K_d - E_0 - S_0)^2 - 4E_0S_0}}{2} \tag{5.7}$$

The measured chemical shift (δ_{eff}) for a fast exchange process is dependent on the ratio of free to bound enzyme (equation 5.8).

$$\delta_{eff} = n_{bound} \cdot \delta_{bound} + n_{free} \cdot \delta_{free} \tag{5.8}$$

where $n_{free} + n_{bound} = 1$.

As only the absolute values of the chemical shifts $\Delta\delta$ were used for the graph depicted in Figure 5.4, the chemical shift of the free state can be subtracted from equation 5.8. This results in the simpler description shown in equation 5.9.

$$\Delta\delta_{eff} = \delta_{eff} - \delta_{free} = \frac{[ES]}{E_0} \cdot \delta_{bound} + \frac{E_0 - [ES]}{E_0} \cdot \delta_{free} - \delta_{free} = \frac{[ES]}{E_0} \cdot \Delta\delta_{bound} \tag{5.9}$$

For a realistic fit function one more minor correction has to be added. As the nucleotide is only a weak binder, the real chemical shift of the bound enzyme cannot be reached. In consequence a factor x describing the surcharge in ratio to δ_{bound} is co-fitted (equation 5.10).

$$\Delta\delta_{eff} = \frac{[ES]}{E_0} \cdot (\Delta\delta_{bound} + \Delta\delta_{bound} \cdot x) \tag{5.10}$$

The calculated K_d for the NMR titration experiment was 1.38 mM with a correction factor $x = 0.187$. This signifies that the actual value of Δ_{av} for the bound protein would have been approximately 20 % higher than the experimentally determined one.
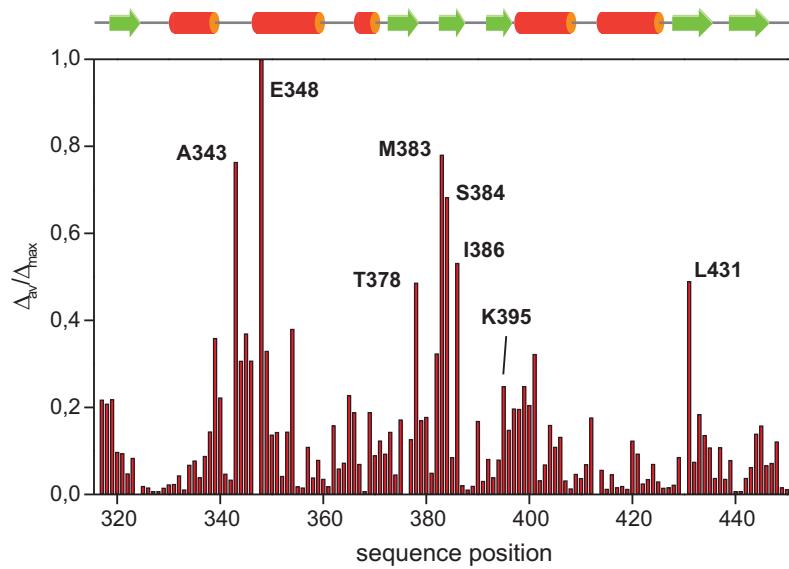
FIGURE 5.5: **Normalised weighted chemical shift differences (Δ_{av}/Δ_{max}) for KdpBN upon AMP-PNP binding per residue.** The secondary structure elements are indicated above. Missing values correspond to the N-terminal residue N316 and the 5 prolines.

The K_d value that was determined by the NMR titration experiment is slightly lower, than the one determined by TNP-ATP displacement, which was calculated to be 2.2 mM. This might be due to the fact that the measurement of fluorescence is much more susceptible to interferences as it is the case for NMR experiments. This can be seen immediately by comparing the smoothness of the graphs. The curves obtained from the TNP-nucleotide displacement experiments are quite bumpy, thus a precise fitting of the K_d value proved to be much more difficult as for the smooth lines of the NMR titration experiment.

The location of the nucleotide binding pocket can be traced by mapping

the differences in chemical shift Δ_{av} on the structure of the protein. Therefore it is useful to normalise Δ_{av} beforehand. This is done by dividing all Δ_{av}'s by the maximal value of the chemical shift difference, which was the one of E348. Figure 5.5 shows the change of Δ_{av}/Δ_{max} along the protein sequence. Strongly influenced are the backbone amide protons and nitrogens situated in the loop region between α-helices 1 and 2 (residues A343 – E348) and the residues on β-strand 3 (M383 – I386).

Mapping of the normalised weighted chemical shift differences ($\Delta_{av}/ \Delta_{max}$) upon the 3D Structure of native KdpBN shows where the nucleotide binding pocket is located in more detail (Figure 5.6). Clearly observable is the concentration of dark coloured residues, corresponding to significantly influenced residues, around the central cleft. Lysine 395, located on β-strand 4, which has previously been determined to be critical for nucleotide binding via FITC labelling, is in the centre of the binding pocket. β-strand 3, which is opposite to β-strand 4 also shows significant chemical shift deviations. The curved β-sheet seems to form one half of the binding pocket. However, the largest changes occur in the unstructured loop region between helix 1 and helix 2, which forms the second half of the binding pocket. The unstructured C-terminus is not affected at all by nucleotide binding, the C- and the N-terminal β-strands are only influenced marginally.

A previous combination of mutational studies and NMR spectroscopy on the Ca^{2+}-ATPase N-domain from rabbit (residues 357–600) identified the nucleotide binding pocket to be at a similar position.[93] Similar to KdpBN they observed low affinity nucleotide binding with an estimated K_d in the range of 10-100 μM. Mapping of the chemical shift deviation Δ_{av}/Δ_{max} of the backbone ^{1}H and ^{15}N amide resonances derived from a series of ^{1}H,^{15}N-HSQC spectra at increasing ratios of nucleotide : protein, showed the most significant changes in regions corresponding to those most affected in the N-domain of KdpB. As only the backbone resonance assignment was accomplished for the Ca^{2+}-ATPase N-domain, a more precise definition of the binding pocket was not made.
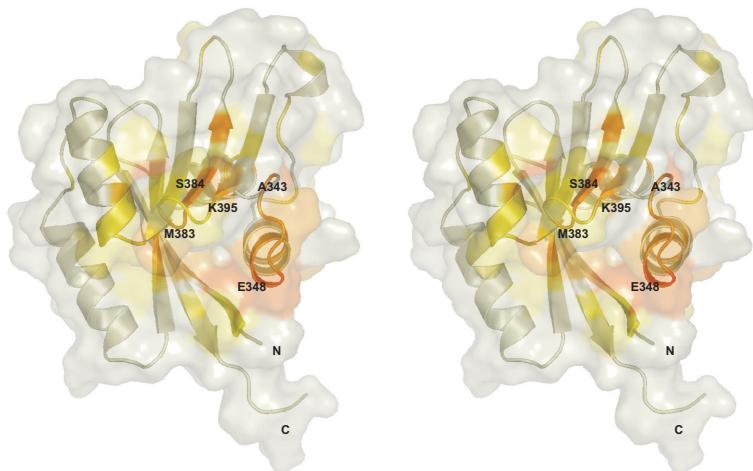
FIGURE 5.6: **Stereo view depicting the mapping of chemical shift changes in the nucleotide binding domain.** Residues are coloured from white ($\Delta_{av}/\Delta_{max} < 10\%$) to orange ($\Delta_{av}/\Delta_{max} > 50\%$). The ones which change most are close to the hydrophobic core, formed by the curved β-sheet and α-helix 2.

5.3 Resonance Assignment of AMP-PNP Bound KdpBN

The basis of further investigations, leading to a more precise definition of the binding pocket, was the resonance assignment of nucleotide bound KdpBN. Therefore, a uniformly ^{13}C,^{15}N-labelled sample of KdpBN was mixed with an 15fold excess of AMP-PNP. The high excess of nucleotide was chosen to ensure an almost complete saturation of the protein throughout the measurements.

As the amide ^{1}H and ^{15}N shifts were already known from the titration experiments, the C^α and C^β resonances could be assigned by applying a CBCA(CO)NH experiment. The H^α chemical shifts and the aliphatic side chain resonances could be assigned to an extent of 97 % by extracting the information provided by an (H)CCH-COSY experiment.
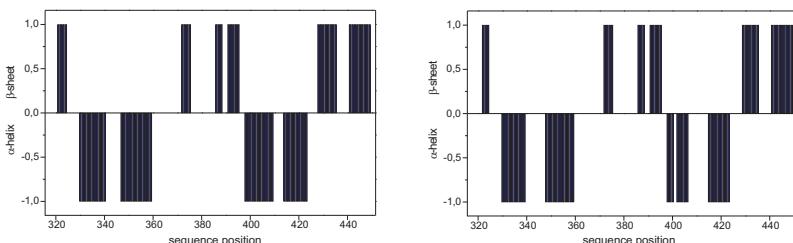
FIGURE 5.7: **Consensus CSI plots of native (left, C', C^α, H^α, C^β) and nucleotide bound KdpBN (right, C^α, H^α, C^β).**

Figure 5.7 shows the evaluation of the secondary chemical shift data, applying the CSI program.[81,82] Though some cross peaks in the ^{1}H,^{15}N-HSQC experiment shift considerably, both consensus plots are almost identical, indicating that no change in secondary structure occurs upon nucleotide binding. Comparison of the side chain chemical shifts of native and nucleotide bound KdpBN only revealed minor differences. Thus, substantial changes in the tertiary structure can also be ruled out.

5.4 Defining the Binding Pocket

5.4.1 The Isotope-Filtered 2D-NOESY Experiment

A good means of identifying intramolecular protein-ligand contacts is the isotope-filtered NOESY experiment.[69] Such an experiment is provided by the ω_2-^{12}C,^{14}N-filtered 2D ^{13}C(H)-NOESY experiment with WATER-GATE.[94, 95] To carry out the experiment, a uniformly ^{13}C,^{15}N-labelled sample of KdpBN was mixed with a 15fold excess of unlabelled AMP-PNP. The pulse sequence starts with an ^{1}H,^{13}C-HSQC step, recording the ^{13}C-chemical shifts of the protein. In this step the stereospecific information gets lost, but this is more than compensated by the higher dispersion and thus increased liability of the cross peaks in the ^{13}C dimension. The HSQC step is followed by a NOESY mixing sequence. Finally, the magnetisation of the ^{13}C and ^{15}N bound protons is transferred into undetectable double quantum coherence. In order to save time and avoid unnecessary losses of magnetisation due to relaxation processes, the ^{13}C,^{15}N-filter was combined with the watergate sequence. The assignment of the resonances of bound AMP-PNP was made with an ^{1}H,^{13}C-HSQC spectrum from a mixture of both unlabelled KdpBN and AMP-PNP. Indeed, the chemical shifts of the aromatic protons H2 and H8 had swapped position compared to the unbound nucleotide.

Figure 5.8 shows the resulting spectrum. The background noise is considerably high. This is plausible, as there is such a huge excess of ^{13}C- and ^{15}N-bound protons that a complete suppression of their magnetisation is not possible, which is partly due to slightly deviating scalar coupling constants and pulse imperfections. Taking a closer look to the spectrum, the 'real' signals are easily distinguished from the background noise. The remaining protein signals have broader line shapes and they usually come in doublets as there is no decoupling during t_2.

The aromatic proton H2 and proton H1' from the ribose ring exhibit a
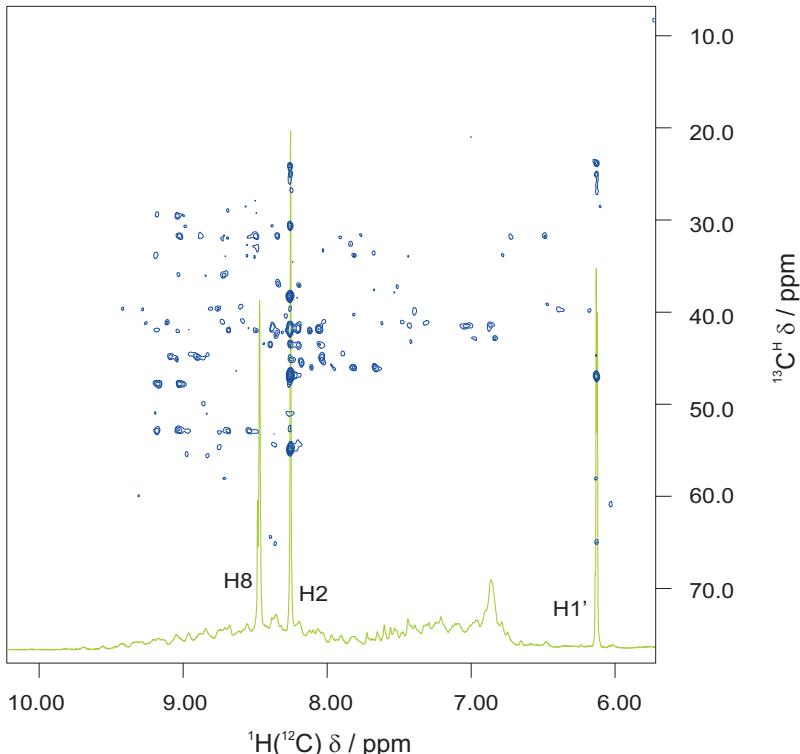
FIGURE 5.8: **The ω_2-^{12}C,^{14}N-filtered 2D ^{13}C(H)-NOESY experiment of uniformly ^{13}C,^{15}N-labelled KdpBN and unlabelled AMP-PNP (ratio 1 : 15).** The 1D spectrum of the protein - AMP-PNP mixture is superimposed (light green). When subtracting the background, it can be seen that only H1′ of the ribose and the aromatic H2 of the nucleotide do have contacts to the protein, the aromatic H8 shows no significant cross peak. The NOE mixing time τ_{mix} was set to 100 ms.

extensive network of NOE cross peaks to the residues K395, S397, M383 and L431. The aromatic proton H8 shows no cross peak, neither do the other ribose-protons (data not shown). H2' is buried under the water signal and can therefore not be observed. The other signals, H3', H4', and both H5' of the ribose are quite close to the water signal and suffer already from the water-suppression, additionally they come as multiplets, which lowers their intensity even more. Further information of the net NOE contacts can be obtained by the STD measurements.

5.4.2 Saturation Transfer Difference Spectroscopy

Group epitope mapping by saturation transfer difference (STD) spectroscopy is a useful means to estimate the amount of contacts a ligand has with the protein it is bound to.[96,97] It is a purely quantitative method, applicable for small, usually non-peptidic ligands. The protein is irradiated with a selective pulse at a resonance where no ligand signals are present. This is preferably the aliphatic region at $+1$ ppm, for large proteins with signals at -2 ppm, it is possible to study peptidic ligands as well. Spin-diffusion quickly leads to an effective saturation of the entire protein, plus the protons of the ligand that are in close contact to the protein surface. The amount of magnetisation transferred from the protein to the single protons of the ligand is dependent on both the proximity to the protein surface and on the number of contacts to protein residues. Prior to acquisition a 50 ms spin-lock pulse was applied ($T_{1\rho}$-filter) to allow the protein to relax. This method efficiently suppresses the signals of the protein, leaving over only the ligand signals. The result can be seen in Figure 5.9. The $T_{1\rho}$-filter slightly decreases the STD signal for the ligand. Therefore a reference spectrum is recorded with exactly the same experimental set-up, only the irradiation pulse is set far outside the spectral window of the protein to $+30$ ppm. This yields a spectrum of the ligand only. The integrals of the ligand signals observed in the STD NMR spectrum are normalised with these values and the most
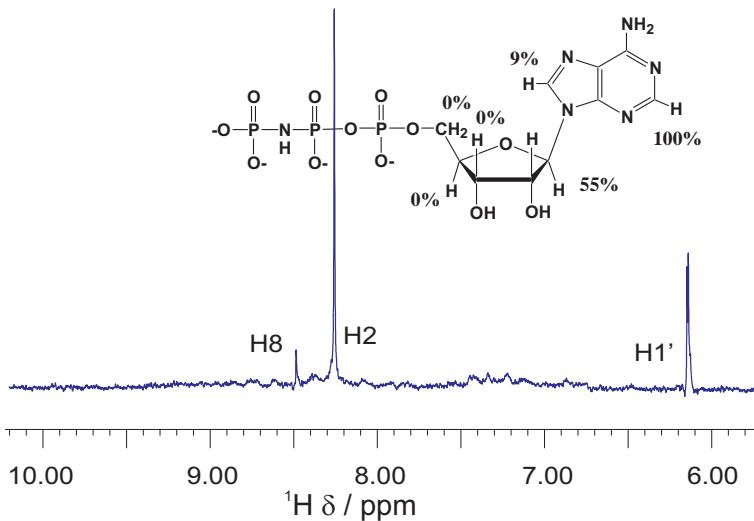
FIGURE 5.9: **Saturation transfer difference experiment on AMP-PNP bound KdpBN.** The strongest transfer can be observed for H2 (100 %), H1' comes in with 55 % whereas H8 only shows 9 % transfer. H3', H4' and H5' don't show any transfer at all. H2' cannot be observed as it overlaps with the water signal (data not shown).

intense signal is set to 100 %. According to that, the relative degree of saturation is determined for the other ligand signals. In the KdpBN – AMP-PNP system, the aromatic proton H2 shows the most efficient transfer of magnetisation. The H1' proton of the ribose sugar follows with a relative degree of saturation of 55 %. H8 only comes in with 9 %. The remaining ribose protons get no detectable amount of magnetisation transferred at all. This result is in accordance with the previous observations from the isotope-filtered NOESY experiment.

These data clearly show that only the aromatic head group is positioned within the binding pocket, whereas the hydrophilic sugar ring is not in close contact to any side chains of the protein.

5.5 The Model of Nucleotide Bound KdpBN

The results obtained from the STD NMR experiment and the ω_2-^{12}C,^{14}N-filtered 2D ^{13}C(H)-NOESY experiment allowed the first calculation of an AMP-PNP bound model of KdpBN. For the calculation 14 intermolecular distance restraints were derived from the isotope-filtered NOESY spectrum. The intramolecular restraints for the protein that were used for the calculation were the same as those used for the free form. Only the restraints to side chains directly within the binding pocket were removed to allow more flexibility and the insertion of the ligand. The use of the same restraints for the calculations seemed non-critical as the evaluation of the chemical shifts of the nucleotide bound protein had previously shown that neither the secondary structure changes, nor the side chain resonances suffer significant changes (see section 5.3). In addition no severe violations of the NOE based restraints occurred when placing AMP-PNP in the binding pocket. This is another hint that no major structural rearrangements within the protein are necessary to accomplish binding.

Figure 5.10 shows the binding pocket in detail. F377 undergoes an aromatic π-stacking with the purine ring system, whilst the ζ-amino group of K395 is drawn close to purine from the other side, presumably forming a cation-π interaction with the purine ring. The purine ring is sandwiched between the side chain residues of strand β 3 and strand β 4 and capped by helices α 1 and α 2 in such way that H8 of AMP-PNP points out of the binding pocket. This explains why no NOE contacts could be observed for this proton. The protons of the ribose ring, except H1', point away from the protein surface into the empty space of the mouth of the pocket. The triphosphate end group protrudes into the solvent via the opening close to the N- and C-termini. It is not stabilised by any side chain contacts to the protein, from this it can be concluded that it is positioned ideally to overcome the distance to the phosphorylation domain.
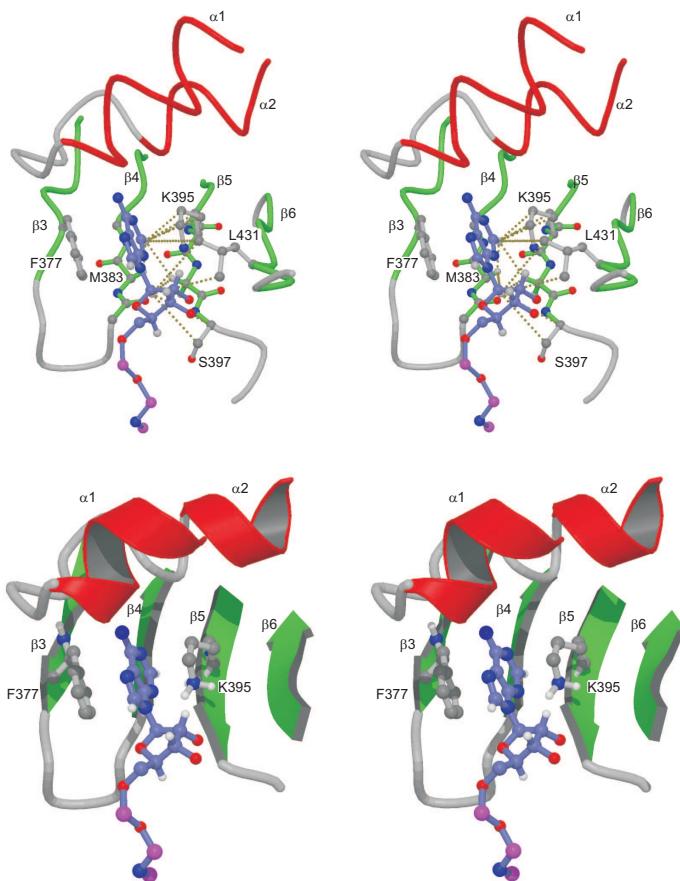
FIGURE 5.10: **The nucleotide binding mode of KdpBN.** The upper panel shows a stereoview of the network of intermolecular NOE connectivities extracted from the filtered 2D NOESY experiment. The lower panel shows the nucleotide to be 'clipped' into the binding pocket, undergoing π-stacking with F377 on one side and forming a cation-π interaction with the positively charged ζ-amino group of K395 on the other.

Chapter

6

Methods and Experiments

6.1 Protein Expression and Isotope Labelling

The corresponding gene fragment of residues Asn[316]–Gly[451] of KdpB was amplified using polymerase chain reaction using the primer pair KdpBN for: 5′–ACACATATGAACCGTCAGGCGTCGGAG–3′ and KdpBN rev: 5′–GCGCTCGAGCTATTAGCCTTTGACGATATCTTTCAG–3′. Primers were purchased from MWG-Biotech (Ebersbach, Germany). Resulting PCR products were cloned blunt into a *Sma*I opened pUC18 vector and transformed into DH5α cells. The sequences were analysed by circle sequencing. Proven constructs were digested with *BamH*I and *Xho*I and the corresponding fragment was cloned into a pET16b vector (Novagen, Groningen, Netherlands). The resulting plasmid pET16bKdpBN was transformed into $CaCl_2$-competent BL21(DE) / pLysS cells (Novagen). *E. coli* BL21(DE) / pLysS / pET16bKdpBN cells were routinely grown in K115 minimal medium (as described by Epstein and Kim, 1971).[98] In order to obtain [15]N- and ^{13}C,^{15}N-labelled proteins for NMR analysis ^{15}N-ammonium sulfate (1.13 g l^{-1}) as nitrogen source and [13]C_{1-6} labelled β-D-glucose as carbon source (2 g l^{-1})

were used. The stable isotopes were purchased from Spectra Stable Isotopes (Columbia, USA). Supplements were added in the following final concentrations: ampicillin, $100\,\mu\mathrm{g\,mL^{-1}}$, chloramphenicol, $60\,\mu\mathrm{g\,mL^{-1}}$, and thiamine, $1\,\mathrm{mg\,mL^{-1}}$. Induction was achieved with 1 mM IPTG when the culture reached an optical density of 0.7. Three hours after induction, cells were harvested, shock-frozen in liquid nitrogen, and stored overnight at $-80°\mathrm{C}$. Pellets were resuspended in 20 mM Tris-HCl pH 7.8, 500 mM NaCl, and 5 mM imidazole. Cells were lysed using a Ribi cell fractionator. Cell debris and membranes were removed by ultracentrifugation at 100.000 g for 60 minutes. The supernatant was applied to preequilibated Ni-NTA (Qiagen, Hilden, Germany) and allowed to bind for 60 minutes on ice under constant shaking. Subsequently, Ni-NTA resign was transferred into a glass column (diameter 1.2 cm) and connected to a FPLC device (Amersham-Pharmacia Biotech, Freiburg, Germany). Final elution was achieved by increasing the imidazole concentration to 125 mM. The Peak fractions were pooled and applied to size exclusion chromatography using a Superdex 75 clolumn (Amersham-Pharmacia Biotech). The column was run with 50 mM potassium phosphate buffer pH 6.0 and 100 mM NaCl. Peak fractions were pooled and concentrated to 1 mM using Centricon micro-concentrators as recommended by the supplier (Millipore, Eschborn, Germany). 0.05 % $\mathrm{Na^+}$-azide were added to prevent microbial growth.

6.2 NMR spectroscopy

The native and the AMP-PNP-bound NMR samples used for structure determination contained 1.4 or 1.0 mM protein, respectively, to which 10 % D_2O were added. The AMP-PNP-bound samples contained a 15 fold excess of AMP-PNP. All spectra were recorded at 300 K either on Bruker DMX600, Bruker DMX750 or Bruker DMX900 spectrometers. Sequential assignment of the native protein was achieved using information on the intra-residue and sequential C^α, C^β, H^α and H^β chemical shifts taken

from CBCA(CO)NH, HNCA, HNCACB, HBHA(CO)NH, HACACO and HN(CA)HA experiments on the double labelled sample and from HNHA and HNHB experiments on a ^{15}N-labelled sample. Additionally, carbonyl carbon assignments were available from an HNCO experiment. Side chain assignment was done using an (H)CCH-COSY experiment. The assignment of ^{15}N and ^{1}H^{N} resonances of the AMP-PNP-bound form was achieved by titration of a ^{15}N-labelled sample with a 94 mM solution of AMP-PNP in 50 mM potassium phosphate buffer at pH 6.0. AMP-PNP was added until complete saturation of the protein had been reached, i.e. no further alteration in the chemical shifts was observable in the ^{1}H,^{15}N-HSQC. The backbone and side chain assignment was done by analysis of a CBCA(CO)NH combined with a (H)CCH-COSY experiment. All NMR data were processed using X-WINNMR and analysed with AURELIA (both Bruker Analytik GmbH, Germany). ^{1}H resonances were referenced to internal 4,4-dimethyl-4-silapentane-1-sulfate (DSS), and ^{13}C and ^{15}N resonances were indirectly referenced to DSS.[83]

Distance data were derived from 3D-NOESY spectra which were all measured with an 80 ms NOE-mixing time. A 3D-^{15}N-HSQC-NOESY and a 3D-^{15}N-HSQC-NOESY-^{15}N-HSQC were recorded on the ^{15}N-labelled sample, and a 3D-^{13}C-HSQC-NOESY, a CCH-NOESY and a CNH-NOESY on the double labelled sample of the native protein.[73] A 3D-^{13}C-HSQC-NOESY, a CCH-NOESY and a CNH-NOESY were recorded on the doubled labelled sample of the AMP-PNP-bound protein as well as an ω_2-^{12}C,^{14}N-filtered 2D ^{13}C(H)-NOESY experiment with WATERGATE to trace intermolecular constraints.[94]

6.3 Structure calculations

Structures were calculated with X-PLOR (version 2.9.3) using standard protocols.[99] NOE cross peaks in the three-dimensional spectra were converted into distance ranges after rescaling according to corresponding HSQC intensities. Cross peaks were divided into four classes, strong, medium, weak and very weak, which resulted in restraints on upper distances of 2.7, 3.2, 4.0 and 5.0 Å, respectively, or 2.7, 3.2, 4.2 and 5.5 Å, respectively, for data acquired at 900 MHz. Lower distance restraints were also included for weak or absent sequential H^N–H^N cross peaks using a minimum distance of 3.2 Å and medium intensity or weaker sequential and intraresidue H^N–H^α cross peaks using a minimum distance of 2.7 Å. Allowances for the use of pseudo-atoms (using r^{-6} averaging) were added where necessary. Dihedral angle restraints for the backbone ϕ and ψ angles were derived from C^α, C^β and H^α chemical shifts using the program TALOS.[100] Restraints were applied for the 92 high-confidence predictions found by the program using the calculated range $\pm 10°$. In addition, direct coupling constant restraints were applied for $^3J(H^N,H^\alpha)$ values measured from an HNHA experiment. Dihedral angle restraints were also applied for the side chain $\chi 1$ angles of the 59 residues where the predominant rotamer was identified during the process of stereospecific assignment. A tolerance of $\pm 45°$ was used with the exception of proline residues where the predominant ring pucker was identified from the $\chi 1$ angle and a restraint of $20(\pm 20°)$ or $-20(\pm 20°)$ applied as appropriate. Hydrogen bond restraints were applied for 57 residues in secondary structure where donor-acceptor pairs were consistently identified in preliminary calculations. The restraints were applied via inclusion of pseudo-covalent bonds as described by Truffault *et al.*.[94]

The structure of the AMP-PNP bound form was calculated using 14 intermolecular distance restraints derived from the 2D ^{13}C(H)-NOESY spectrum. Restraints for the protein were those used for the free form, with the removal of restraints to side chains directly within the binding pocket. Three

additional restraints were used to impart coplanarity to the aromatic rings of the ligand and F377.

6.4 Extent of Assignment

The full assignment of all possible $^1H^N$ and ^{15}N resonances could be accomplished for both, the free protein and the AMP-PNP-bound form. All C', C^α, H^α and aliphatic side chain resonances were assigned in the free protein. For the AMP-PNP-bound form assignment of 99% of the C^α, H^α, C^β, H^β resonances and 95 % of the remaining aliphatic side chain resonances could be completed.

Secondary structure elements were determined by the analysis of secondary chemical shifts for C^α, H^α, C^β and C' in case of the free protein and for C^α, H^α and C^β in case of the AMP-PNP-bound form. The consensus chemical shift index (CSI) plot was obtained using the program of Wishart & Sykes.[80–82]

6.5 Data Deposition

The chemical shifts for the native and the AMP-PNP-bound N-domain of KdpB have been deposited in the BioMagResBank database under the BMRB accession number 6029 and 6030, respectively.[79]

The coordinates for the structure ensemble and a regularised average structure for the native form of KdpBN have been deposited in the Protein Data Bank (accession codes 1SVJ and 2XYZ, respectively). A representative structure for the AMP-PNP bound form has also been deposited (3XYZ).[34]

6.6 Artwork

The three dimensional pictures of proteins were either produced with Py-MOL[101] or a combination of MOLSCRIPT[102] and RASTER3D.[103] Chem-

ical formula of small molecules or sections of a protein chain were drawn
with MDL Isis/Draw.[104]

Chapter

7

Discussion

7.1 The 3D Structure of KdpBN and Nucleotide Binding

The nucleotide binding domain of KdpB proved to be well structured and to exhibit a high degree of internal symmetry, concerning its structural properties. It consists of a curved, six-stranded, antiparallel β-sheet with two flanking α-helices on either side of the β-sheet. Only short unstructured loops connect the secondary structured elements.

In the meantime, when the investigations on KdpBN were accomplished, the NMR-structure of the nucleotide binding domain of the Na^+,K^+-ATPase from rat, at type IIC P-type ATPase, was published. The structure was solved at 0.95 Å resolution in the nucleotide free state (PDB accession code 1MO7) and at 1.0 Å resolution in the nucleotide bound state (1MO8).[105] Shortly afterwards the crystal structure of the N-domain of a closely related Na^+,K^+-ATPase from pig was published by Håkansson at 2.6 Å resolution (1Q3I).[106] Comparison of KdpBN to the N-domain structures of the Na^+,K^+-ATPase crystal structure (1Q3I)[106] and solution structure (1MO7)[105] and the Ca^{2+}-ATPase crystal structure (1EUL)[23] shows KdpBN
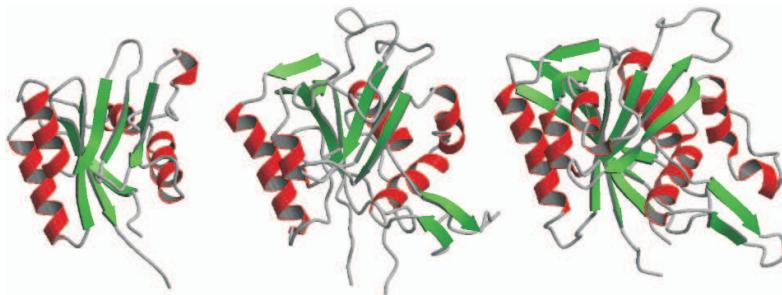
FIGURE 7.1: **Comparison of the structures of the nucleotide binding domains of P-type ATPases.** KdpBN, the Na⁺,K⁺-ATPase (1Q3I)[106] and the Ca²⁺-ATPase (1EUL)[23] are shown from left to right. KdpBN represents the basic core motive, a six-stranded, antiparallel β-sheet flanked by four α-helices. Structural alignment of the backbone of the core motive of the N-domains yielded an RMSD of 2.7 Å for KdpBN vs. 1Q3I and of 2.8 Å vs. 1EUL.

to represent the core N-domain motive (Figure 7.1). Pairwise structural alignments (using DALI[107]) yielded RMSD's of 2.7 Å and 3.2 Å for KdpBN versus 1Q3I and 1MO7, respectively, and 2.8 Å for 1EUL, which indicates good structural agreement in the core region. The structures of the other N-domains possess additional α-helices, an extended β-sheet and two additional β-hairpins, one joining the β-sheet at the top of the N-domain, the other protruding into the solvent and pointing towards the phosphorylation domain.

The additional α-helices and β-strands apparently influence the nucleotide binding mode of the individual domains. For the nucleotide bound Na⁺,K⁺-ATPase N-domain (1MO8) it was reported that the ribose ring is firmly anchored to the protein via several protein-sugar NOE contacts,[105] whereas no NOE cross peaks from the protein to the sugar moiety were observed for KdpBN, except for proton H1′. Thus it seems that the KdpB N-domain displays not only a minimal scheme for the nucleotide binding

domain, but also a rudimentary ATP binding motive comprised of F377, K395 and G396. These residues are completely conserved among the type II-IV P-type ATPases (see Figure 2.7, section 2.3). It might be speculated that additional interactions with ATP found in the Na^+,K^+-ATPase and Ca^{2+}-ATPases are secondary acquirements which resulted in higher ATP binding affinities. The ATP affinity in case of the Ca^{2+}-ATPase is 30 μM,[108] compared to 80-120 μM in case of KdpB.[109] Therefore it is most likely that F377 and K395 provide the basis for nucleotide binding, whereas G396 seems only to be conserved for steric reasons (Figure 7.3 and Figure 5.10, section 5.5). The proton $H^{\alpha 2}$ of glycine points directly into the binding pocket. Any other amino acid in this place would hinder the incoming ligand with its side chain from being placed correctly on site.

The binding mode presented here shows the aromatic ring of AMP-PNP to be *clipped* into the binding pocket by π-interactions on both sides. On one side phenylalanine 377 undergoes an aromatic π-π-stacking with the purine ring from the nucleotide. A classic π-π-stacking between phenylalanine and purine can be estimated to contribute a binding energy of approximately -2 kcal/mol.[110] Lysine 395 stabilises the complex via a cation-π interaction by approaching the purine ring with the positively charged ζ-amino group from the other side. Recent studies have shown that cation-π interactions can provide strong and specific interactions in proteins, especially when exposed to the solvent or when water is displaced from a binding pocket. In aqueous solutions the magnitude of cation-π interaction energy is estimated to contribute about -5.5 kcal/mol.[111] An overall binding energy of approximately -7.5 kcal/mol is strictly speaking not much, but considering the low dissociation constant K_d of 1.4 mM for the nucleotide binding domain alone and ~0.1 mM for the complete protein with the second binding site for the triphosphate end-group present at the phosphorylation domain, the energy summoned up is definitely sufficient.

The nucleotide binding mode of KdpBN differs significantly from other ATP-binding domains. A very well examined group are the ABC transport
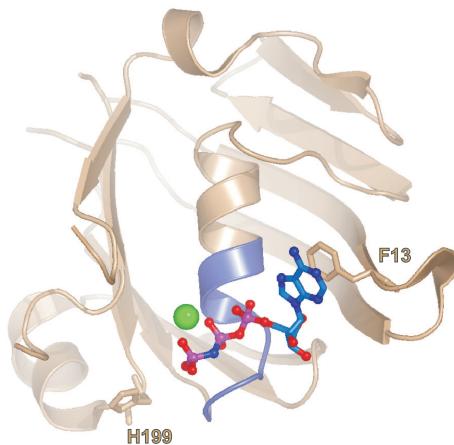
FIGURE 7.2: **Nucleotide binding mode of P-loop ATPases.**[113] The purine ring is attached to the protein surface via an aromatic π-stacking interaction with F13. The central helix unwinds by one turn to enclose the α- and β-phosphate groups with the backbone amide protons. The P-loop (or Walker A motive) is coloured in purple. A Mg^{2+} ion, stabilised by side chain contacts, ushers the γ-phosphate group to stretch towards H199, the site of phosphorylation.

proteins, which belong to the P-loop ATPase superfamily. These enzymes derive considerable binding energy from the triphosphate group of ATP capping the N-terminus of a helix by introducing the α- and β-phosphate group in the first turn of this helix. In addition, a hydrophobic π-π-interaction of the nucleotide with a conserved phenylalanine or tyrosine residue on the surface of the protein accomplishes the binding (Figure 7.2).[112, 113] This results in a micromolar binding constant of approximately 118 μM for ATP and 139 μM for ADP.[114] In contrast, the current investigations on KdpBN show that only the aromatic purine system is essential for ATP binding in the P-type ATPase N-domains, whilst the triphosphate end group projects into the solvent (Figure 7.3).
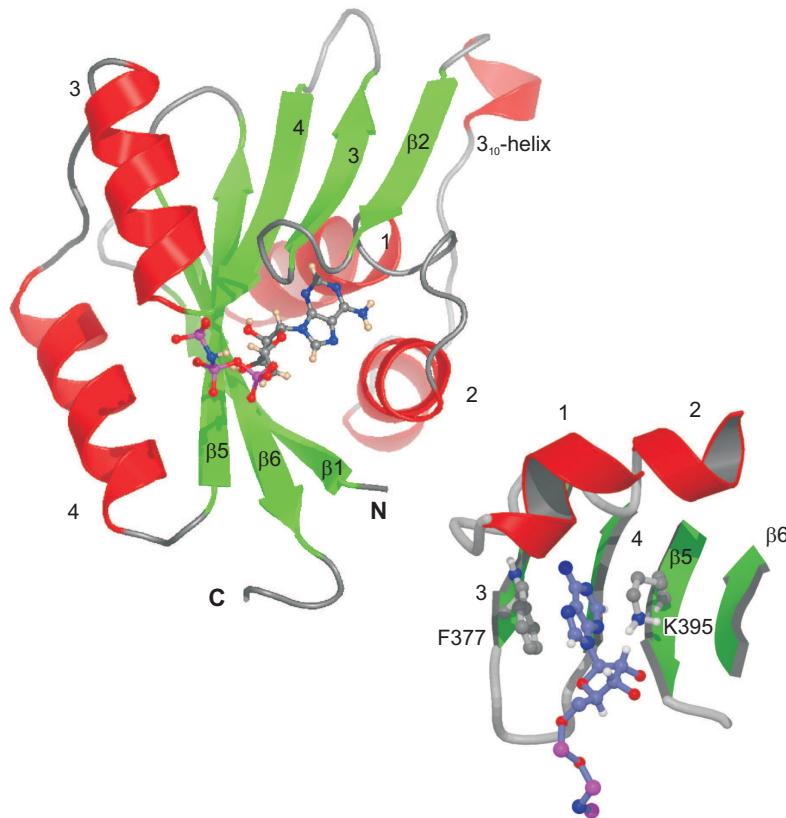
FIGURE 7.3: **The model of nucleotide bound KdpBN.** Left, the whole N-domain is depicted. The aromatic part of the nucleotide is buried in the hydrophobic core of the domain, whilst the hydrophilic triphosphate end-group protrudes into the solvent. Right, the binding pocket is zoomed out. The *clipping* mechanism (including F377 – AMP-PNP – K395) can clearly be observed.

7.2 The Reaction Cycle of the KdpFABC Complex

The dual function, ion transport and ATP hydrolysis, of the central subunit of P-type ATPases, with binding of two or more substrates, reflects a complicated reaction cycle which has been simplified in the so called Post-Albers scheme (see Figure 2.8, section 2.4).[44,45] In this scheme the enzyme has two main states; E1, with low-affinity ion binding sites on the release side of the membrane and E2, with high-affinity ion binding sites on the uptake side. Conformational changes in the transmembrane domain and rigid body re-arrangement of the cytoplasmic domains during the E1 / E2 transition was persuasively demonstrated by the structures of the Ca^{2+}-ATPase at different stages of the cycle.[23,53] Only recently the crystallisation of the Ca^{2+}-ATPase was achieved in presence of the nonhydrolysable ATP analogue AMP-PCP independently by the groups of Møller and Nissen[115] at 2.6 Å resolution and Toyoshima at 2.9 Å resolution.[116] Both structures show the transmembrane helices in the arrangement typical for the E1 state. In contrast, the cytoplasmic domains are arranged similar to the thapsigargin-bound E2 state.[53] This is not surprising, as the rigid body translocation of the cytoplasmic domains has been recognised earlier as a prerequisite to bring ATP close to the site of phosphorylation before the transmission of the γ-phosphate group to the conserved aspartate residue can take place. The hydrolysation energy then leads to a rearrangement of the transmembrane helices and pumping of the corresponding ions. However, to understand the energy transduction process in this type of enzyme, the mechanism of the cytosolic rigid body movement has to be elucidated.

One notable feature of the Ca^{2+}-ATPase structure in the presence of bound Ca^{2+} is that the N-domain is well separated from the phosphorylation site. In order to deliver ATP to this site, the N-domain must undergo a considerable rigid body movement. In the solution structure of the Na^+,K^+-ATPase N-domain, Hilge *et al.* observed a conformational change upon ATP binding, which they interpreted to be the driving force for this interdomain rear-

rangement.[105] This change was centred on S583 in the C-terminal β-strand, which was reported to undergo a flip from phi/psi angles of $164°/57°$ in the free form to $-79°/178°$ in the bound form, leading to a repositioning of the hinge region between the N- and P-domains. A key result of the current investigations is that this conformational change has not been observed for the KdpB N-domain. The reason for this difference is a discrepancy in the conformations of the free forms of the domains. The residue in KdpBN corresponding to S583, A444, has phi/psi angles of $-91.6°/118.9°$ and thus has a conformation more closely resembling that observed by Hilge *et al.* for the nucleotide bound form. Thus the possibility that KdpBN undergoes an equivalent conformational change can be excluded. In fact, chemical shifts and NOE connectivities for residues in the C- and N-terminal strands show no significant changes upon addition of AMP-PNP, indicating that the local conformation remains unchanged in this region. Figure 7.4 shows a strip from the CNH-NOESY spectrum of the NOE cross peak pattern of alanine 444 for the free and the nucleotide bound form of KdpBN.

Although this could represent a real difference between the two proteins, the resolution of the Na^+,K^+-ATPase N-domain structure is rather poor in this region and the secondary structure does not conform with that observed for KdpBN or in all other available crystal structures. It is possible that the secondary structure has been disrupted as an artefact of isolation of the N-domain or of the high pH used (8.6). In this case, the conformational change observed has no physiological relevance. Furthermore, the low ATP binding affinity of the N-domain determined both biochemically and via NMR titration experiments, signifies a low binding energy and fast exchange between bound and unbound forms, making major structural rearrangements unlikely.

The absence of an ATP-induced conformational change in the N-domain supports the idea that the N-domain is flexibly linked to the P-domain and that it engages and disengages the P-domain at various stages of the reaction cycle through thermal motion. A crystal structure always represents
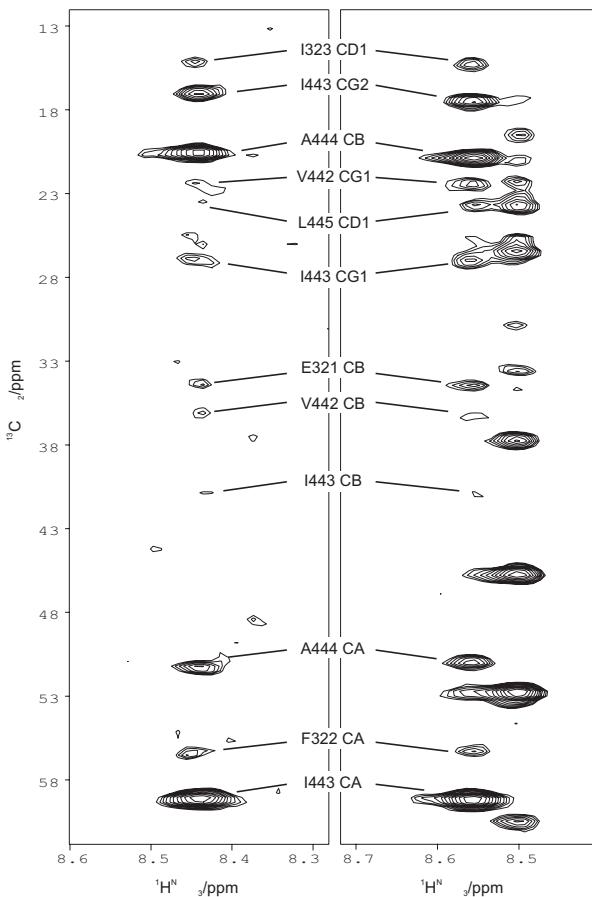
FIGURE 7.4: **Strips from the CNH-NOESY experiment.** Shown are the strips for
the amide proton of A444, which can be structurally aligned to S583 in the NMR-
structure of the Na$^+$,K$^+$-ATPase of Hilge *et al.* (1MO7).[105] The left strip is from
the nucleotide bound sample, recorded at 750 MHz, the strip on the right is the free
protein at 900 MHz. An identical pattern of NOE contacts is clearly visible in both
strips, indicating that no conformational change occurs in the region of A444 upon
nucleotide binding. The assignment displayed is this of native KdpBN.

only a snapshot of the enzyme conformation at a certain stage. The lattice of the crystal can create local energy minima that are not present when the protein is in its native environment. The very open structure displayed by the Ca^{2+}-ATPase in the Ca^{2+}-bound form probably represents one extreme of the N-domain's flexible range, stabilised by crystal contacts, and therefore needs not represent the range covered during the reaction cycle. In contrast, it is more likely that the ATPases do not necessarily undergo such severe rigid body motions of an angle of $110°$ as observed for the N-domain of the Ca^{2+}-ATPase at each turn of the reaction cycle as this would represent a very time and energy consuming procedure which would prevent the fast turnover rate of the cycle. It can be hypothesised that the actual rigid body movements of the cytoplasmic domains during the reaction cycle are only at a scale of minimal vibrations, just large enough to allow an effective exchange of ADP by ATP and docking of the A-domain to the P-domain to hydrolyse the phosphorylated aspartate residue and re-establish activity. The results of the current study thus extend the view that the cytoplasmic domains of P-type ATPases show only minor internal conformational changes upon ligand binding, but are free to undergo considerable rigid body movement with respect to each other.

The concept of a flexible N-domain has important implications for the reaction cycle (Figure 7.5). It is commonly accepted that binding of the transported ion is a key step leading to a high ATP affinity conformation (E1 state). A flexibly linked N-domain implies that this increase in ATP affinity is the result of activation of the phosphorylation site rather than any effect on the N-domain itself, i. e. the high-affinity ATP binding site is a combination of the N- and P-domain sites formed by engagement of the N-domain. For the reaction cycle this signifies that as long as there are no ions to be transported, the N-domain swings back and forth with cytoplasmic ATP being bound and released presumably with a dissociation constant K_d that is similar to that of the isolated N-domain, KdpBN (E2 state). Binding of an ion causes the rearrangement of the transmembrane helices, here sym-

bolised by stretching of the TM-helix 5, leading to the E1 state. For the unlikely case that no cytoplasmic ATP is present, the N-domain is still free to swing back an forth. This is with high probability the scenario which led to the open conformation in the crystal structure of the Ca^{2+}-ATPase in the Ca^{2+}-bound state (1EUL). Binding of the ion activates at the same time the phosphorylation site in the P-domain. If now the N-domain approaches the P-domain with an ATP bound, it is drawn towards it. The presence of two binding sites for the nucleotide increase the affinity for ATP drastically. In case of the Kdp system the apparent K_d at pH 7.8 rises from 0.77 mM for the isolated N-domain (KdpBN) to 100 μM for the intact KdpB protein.[109] The affinities of two binding sites multiply if they are subjected to the same ligand. This corresponds to a dissociation constant of approximately 0.1 M for the phosphorylation site, a value that is indeed low and explains the necessity of an additional nucleotide binding domain. One great advantage of a twofold nucleotide binding site is that only minor changes in affinity in one site lead to a considerable change in the product of both affinities. This feature compensates for the increased complexity of the protein as it facilitates and speeds up the switching process from high to low affinity complexes. Transition state I (TSI) was introduced in the Post-Albers-scheme to describe this complex interaction of the domains. From the data available, no evidence for the position of the A-domain at the transition state can be derived. Presumably it is at this very moment turning the tip containing the TGES motive towards the P-domain, as indicated by the circular arrow. As the whole system is in continuous flow, the precise definition of the enzyme at a definite stage seems neither possible nor sensible.

Nucleotide binding and subsequent activation of the γ-phosphate group leads to the E1~P state. The formation of the phosphoenzyme causes a conformational change of the transmembrane region which lowers the ion affinity (E2-P state). The transported ions are released into the cytoplasm. For this stage of the reaction cycle an atomic model of the Ca^{2+}-ATPase from a 6 Å structure by cryoelectron microscopy exists (1KJU).[117] The crystals
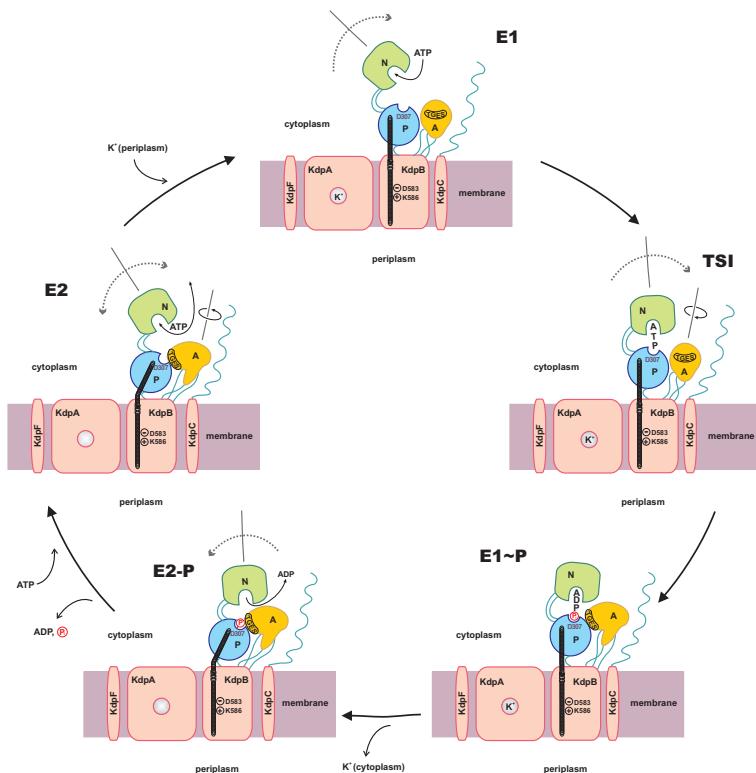
FIGURE 7.5: **The proposed reaction cycle for the KdpFABC complex.** The E2 state has high ion affinity on the extracellular side of the membrane. ATP, due to its low binding constant, is in rapid exchange with the flexibly hinged N-domain (movement indicated by arrows). K^+ binding triggers the start of the reaction cycle, leading to the E1 state with increased ATP affinity. The phosphorylation site (D307) is activated and the N-domain engages; the nucleotide is now bound between the N- and P-domains (TSI). Phosphorylation takes place, the A-domain engages via the TGES motive (E1~P state) followed by conformational changes in the transmembrane domain leading to the E2-P state with low ion binding affinity on the cytoplasmic side of the membrane (release of K^+). Subsequently D307 is dephosphorylated by the A-domain (E2 state). Now periplasmic K^+ binds, the A-domain disengages and KdpB returns to the E1 state.

were formed in presence of decavanadate and thapsigargin, but in absence of Ca^{2+}. It is well established that thapsigargin blocks P-type ATPases in the E2 conformation and that decavanadate attaches to the phosphorylation site and thus mimics the phosphoenzyme. The model of the Ca^{2+}-ATPase in the E2-P state shows the N-domain still in contact with the P- and A-domains. This again might represent only one possible conformation which was favoured in the crystal lattice. It is imaginable that as soon as ATP, displaying the connecting link between the domains, is no longer present, the N-domain is at this point already disengaging, deliberating ADP from the binding pocket. The A-domain is now free to stimulate dephosphorylation and the enzyme returns in the E2 state where the reaction cycle continues anew with the uptake of an extracytoplasmic ion.

This model of the reaction cycle is consistent with the observation that ATPase activity is coupled to ion binding and not vice versa, not only in the Kdp-ATPase but also in all other P-type ATPases. The conclusions drawn here are summarised in Figure 7.5.

The impulse to discuss the traditional E1E2 model under new light has been encouraged only recently by Scarborough.[48, 118, 119] He strongly bemoans the restriction to only two main-conformations, E1 and E2, and the resulting oversimplification which may lead to a misinterpretation of the orientation and affinities of the ion-binding sites. This is indeed an ambiguous subject. E1 has per definition of the classical E1E2 model high ion-affinity on the uptake side of the membrane.[46] This hypothesis is no longer serve as a base for the insights derived from the crystal structures of the Ca^{2+}-ATPase. Those have shown that E1 only exists in an ion-bound form and that the formation of the E1 state is driven by ion-binding to the E2 conformation, whereas the E1E2 model implies a spontaneous transition of E2 to E1 before binding of the extracytoplasmic ions takes place. Scarborough therefore suggests to abandon readily the E1E2 nomenclature and to apply a five step reaction cycle that is not defined by affinities but rather by intermediate states and transition states. One inherent advantage of the E1E2 nomenclature is
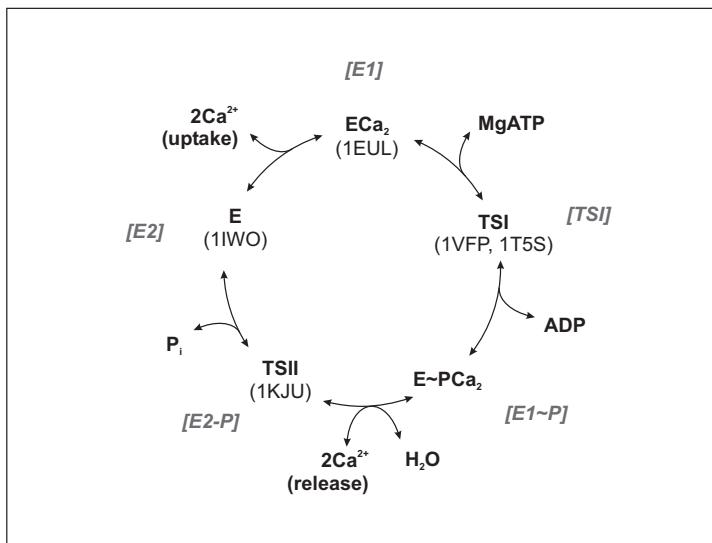
FIGURE 7.6: **The reaction cycle of P-type ATPases according to Scarborough.**[118] Depicted is the 'new' nomenclature according to Scarborough and if available the PDB accession code for the crystal structure. The nomenclature of the individual steps as applied for the Kdp ATPase in Figure 7.5 is given along in square brackets.

the fact that the known crystal structures have been assigned to the individual states. Thus it is only necessary to get rid of the assumption of the predefined affinities. Figure 7.6 depicts the suggestions of Scarborough and the relation of the individual states to the reaction cycle of the Kdp ATPase as shown in Figure 7.5.

The recently published crystal structures of the Ca^{2+}-ATPase with the nonhydrolysable ATP analogue AMP-PCP bound, 1VFP[116] and 1T5S,[115] seem to represent the postulated transition state I. Their conformation is in good agreement with the predictions that were made for the Kdp ATPase reaction cycle. Both structures do not reflect the conformational change in the

C-terminal β-strand of the N-domain as observed by Hilge *et al.*,[105] which strongly supports the results of the current study that ATP does not induce inter-domain motions, but stabilises the engagement of the nucleotide binding domain and the phosphorylation domain.

When discussing the function of the Kdp ATPase in relation to other P-type ATPases the singular architecture of the Kdp system has to be carefully considered. Theoretically, all steps in the reaction cycle of P-type ATPases are reversible, thus they could function under appropriate circumstances as ATP-synthases.[120] This is no longer true for the KdpFABC complex as transport is not mediated by the P-type ATPase subunit KdpB, but by the KdpA subunit, which is a one-way system due to its similarities to the KcsA channel protein.

7.3 Evolutionary Aspects

Lysine 395 of KdpB has been shown to react with FITC resulting in the inhibition of ATPase activity.[109] The structural data provide clear justification for this inhibition, both in KdpB and all P-type ATPases sharing the KGXXE/D motive. Interestingly, these residues are absent in the heavy metal-transporting P-type ATPases (type IB), rendering the *clipping* mechanism of the aromatic moiety of the nucleotide as observed for KdpBN impossible. Therefore the conclusion can be drawn that nucleotide binding in type IB P-type ATPase differs somewhat from that of all other representatives of the family. Recent extensive mutational studies on the Zn^{2+}-transporting P-type ATPase from *Escherichia coli* suggest an involvement of a conserved glycine motive in nucleotide binding.[121] This gives rise to the assumption that KdpB is currently misgrouped as a class I P-type ATPase.[21,122] Thus the evolutionary tree of P-type ATPases may need to be redrawn, since KdpB seems to bear more similarities to the plasma membrane proton pump (class III), than to the heavy metal-transporting class I P-type ATPases.

The modular design of the P-type ATPases gives rise to the assumption

that at very early stages of evolution the enzyme was constructed of four independently operating proteins; one hydrophobic subunit, spanning the membrane and three individual cytoplasmic proteins, comprising the phosphorylation site for energy transfer, the nucleotide binding pocket and the actuator site catalysing dephosphorylation. At some stage the proteins merged stepwise to increase functionality. Possibly the heavy metal transporting ATPases fused at this time another nucleotide binding protein. Subsequent mutation created substrate specificity. Additionally, KdpB lost the ability of autonomous ion transport, by capping the transmembrane helices 8–10, but captured an ion channel instead, which it regulates now in such way that it works against the ion gradient.

Chapter

8

NMR Structure of an $\alpha 4 \beta 7$ Integrin Antagonist

8.1 Integrins

Higher organisms, consisting of multiple cells, have developed several systems to communicate cell–cell and cell–matrix contacts. The extracellular matrix (ECM) is composed of a complex mixture of macromolecules that are secreted by fibroblasts of the connective tissue. There are two main groups of theses macromolecules, glycosaminoglycans and fibrous proteins. The glycosaminoglycans are polysaccharide chains that exhibit the basis of the extracellular matrix. They form a gel-like three dimensional environment in which the fibrous proteins are embedded. The functionality of fibrous proteins is either of structural or of adhesive character. Structural fibrous proteins such as collagen and elastin shape the extracellular matrix and lend it elasticity. The adhesive fibrous proteins mediate, according to their name, cell–matrix contacts. Fibronectin is one of the most abundant proteins in the extracellular matrix. It forms a dimer that is covalently linked by disul-
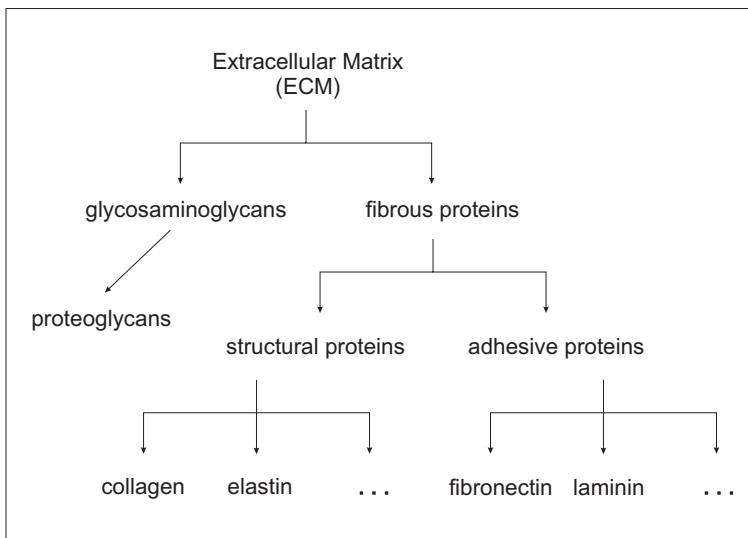
FIGURE 8.1: **Flow chart of the organisation of selected compounds of the extracellular matrix.**

fide bridges. Each monomeric subunit is folded in three functional distinct domains which have been assigned to bind specifically to collagen, heparin and individual integrins, respectively. Figure 8.1 provides a schematic overview of the composition of the extracellular matrix, shown by example of some selected main compounds.

Cell adhesion molecules (CAMs) are proteins at the cell surface that mediate and control cellular behaviour and development by attaching to the adhesive fibrous proteins of the extracellular matrix. They are interleaved in processes as adhesion, migration, maturation, differentiation and metabolic function of their hosting cells. Thus, they are vital for embryogenesis, regenerative processes of tissue, immune response and many more. But they can also have pathogenic properties, as they enhance angiogenesis,

metastasis of tumours, thrombosis and so on. Cell adhesion molecules are grouped into four classes; cadherines, immunoglobulines, selectines and integrins.[123] These molecules discern particular peptide sequences on the adhesive fibrous proteins of the extracellular matrix.

Integrins are the largest and most versatile family of these receptor proteins. They do not only anchor cells to the matrix, but they also function as receptors for distinct matrix proteins. Thus they are involved in *outside-in signalling*, i.e. they transfer environmental stimuli into the cell by binding of an extracellular ligand and *inside-out signalling*, i.e. the affinity of the integrins can be adjusted according to the actual cellular requirements from a passive, slightly adhesive state (responsible for tethering and rolling) to an active, strongly adhesive state.[124, 125] Thereby migration, spreading and maturation of the cells are regulated.

Integrins are heterodimeric transmembrane glycoproteins, consisting of an α and a β subunit. Currently, there are 18 α and 8 β subunits known, which hetero-dimerise to yield 24 distinct combinations.[126] Common to both subdomains are the large extracellular domains, one transmembrane spanning helix and small intracellular domains of only 20–50 amino acid residues. The α subunit is made of approximately 950–1140 amino acid residues. The extracytoplasmic N-terminal domain consists of seven repeat domains that assume a seven bladed β propeller. The head group is linked by an almost linear stalk region to the transmembrane helix. The β subunit is slightly smaller with only 730–800 residues with the exception of $\beta 4$ which has got 1750 amino acid residues. It assumes a Rossmann fold and is anchored to the membrane by an equally long stalk region as found in the α subunit.[127, 128]

The head groups of both subunits are involved equally in ligand-binding. Most integrins do not bind singularly to a specific ligand, but they can bind either to proteins of the extracellular matrix as fibronectin or to receptors on the surface of other cells. The recognition sequence of the ligands are versatile and often an integrin ligand has several different recognition sites

for interaction with different integrins. Fibronectin interacts for instance with $\alpha X\beta 2$ integrins via the GPRP sequence and with $\alpha IIb\beta 3$ via the RGD sequence.[129] Despite of the variety of possible integrin ligands, there exist two most favoured recognition sequences, RGD and LDV. Whereas the RGD sequence is accepted by approximately 50 % of the integrin family, LDV is less common.

The mechanism by which integrins function has given rise to many controversial debates. The currently most favoured mechanism is the three-state model which is depicted in Figure 8.2.[128, 130, 131] This model bases on a global search of similar helix-helix interactions in transmembrane proteins. The transmembrane helices play a major role in this model. This could be confirmed by later recalculation of the model, exploiting the data from an NMR structure of the cytoplasmic domains.[132, 133]
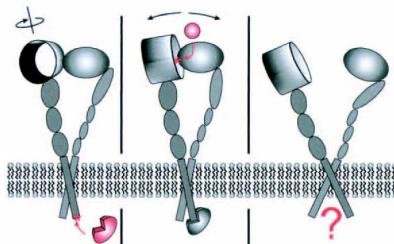
FIGURE 8.2: **Putative three state mechanism of integrins.**   Figure taken from Gottschalk *et al.*.[128]

Initially, the integrin is in a resting position. Binding of a ligand to the cytoplasmic domain of the integrin triggers a rotational movement of the transmembrane helices of approximately 90° in respect to each other. This rotation causes simultaneously a switch in the positions of the head groups, leading to an activated state by exposing the extracellular binding site. Ligands are now free to bind and the head groups are separated by a scissor-like movement. Ligand binding induces the formation of clusters, where

the α and the β subunit homodimerise or homotrimerise, respectively. A recent computational model based on known X-ray and NMR-structures and mutational results, suggest two different models of clustering and signal transduction.[134]

8.2 $\alpha 4\beta 7$ Integrin Antagonists

$\alpha 4\beta 1$, $\alpha 4\beta 7$ and αE$\beta 7$ integrins are mainly expressed in cells of the immune system. They are involved in inflammatory and immune response processes.[135] The spreading of leukocytes includes several steps, *tethering*, *rolling*, *activation* of specific receptors, *adhesion*, *diapedosis* and finally *homing*.[136] The first loose contact, tethering, is mainly mediated by selectines and in some cases by $\alpha 4\beta 7$ integrins. Tethering slows down the rotational velocity of the floating leukocyte. This allows the leukocyte to search the surface of the blood vessel for adhesive factors. Activation of specific receptors upon recognition of specific factors leads to adhesion of the leukocyte. Then the leukocyte penetrates the vascular wall (diapedosis) and is directed towards the particular tissue where it eventually attaches (homing). The spreading of leukocytes into the surrounding tissue is an important immunological process. The control of local immunity, mode and intensity of immune response is managed by the regulation of the permeability of the vascular wall for specialised leukocytes.

$\alpha 4\beta 7$ integrins regulate the homing of $\alpha 4\beta 7$ integrin expressing leukocytes to mucosa associated lymphatic tissue. This is why they are often cited as LPAM-1, lymphocyte Peyer's patch specific adhesion molecule. In grown up humans $\alpha 4\beta 7$ integrins are expressed in varying amounts on T- and B-lymphocytes. Assays on $\beta 7$ knock-out mice indicated that $\beta 7$ integrins are indispensable for the formation and maintenance of the mucosa lymphatic tissue.[137]

$\alpha 4\beta 7$ integrins have shown to be promising targets for treating inflammatory and autoimmune diseases. They bind via a Leu-Asp-Thr sequence

(LDT sequence) specifically and with high affinity to the mucosal addressin cell adhesion molecule 1 (MAdCAM-1).[138, 139] The affinity can be enhanced by activation of the leukocytes. A complex interaction of selectines and $\alpha 4 \beta 7$ integrins controls the distinction of the tissue in question. Thus, the expression rate of $\alpha 4 \beta 7$ integrins on leukocytes influences homing.

8.3 Prerequisites for Drug Design

Drug design is a challenging issue in pharmaceutical chemistry. The way from a bioactive substance over the lead compound to the purchasable product is usually not straightforward. To set standards for a rational and responsible approach to new compounds the ADME parameters were defined. ADME stands for absorption, distribution, metabolism and excretion. Only drugs that meet the international standards for these four basic requirements are considered for further development.

The point that has to be addressed first is the problem of absorption and distribution of a drug within the body. For the most efficient transport of a drug to its point of destination, it has to have both, liphophilic and hydrophilc properties. Consequently it can be solvated in the plasma for the issue of transport and penetrate cell membranes the reach the target. The lipophilic properties of a substance can be defined via its partition coefficient P between 1-octanol and water (equation 8.1) where α is the dissociation constant of the substance in water.[140]

$$P = \frac{[\text{substance}]_{\text{octanol}}}{[\text{substance}]_{\text{water}} \cdot (1 - \alpha)} \tag{8.1}$$

It is common practice to refer to the logarithm of the partition coefficient, log$P$, for an easier classification of the drugs. For substances with an higher solubility in 1-octanol it holds $P > 1$, thus log$P$ is positiv. For an higher solubility in water $P < 0$ and log$P$ gets negative.

There are many other parameters that influence the bioavailability of a

drug. In addition there are further requirements to meet that facilitate the application. Oral available drugs have an higher acceptance among patients and sell better. In order to reduce costs during the highly expensive drug development process, Lipinski *et al.* from *Pfizer* company established a scheme that facilitates the early recognition of promising candidates. Therefore they screened more than 3000 active substances from drug databases for common properties. The excerpt of this extensive search is condensed in the *Pfizer's rule of five*:[141, 142]

1. molecular weight < 500 g/mol

2. $4 \leq \log P \leq 5$

3. $\sum$ hydrogen bond acceptors (N, O) < 10

4. $\sum$ hydrogen bond donors (NH, OH) < 5

5. one rule may be violated

A common counter-example to these rules is the successfully applied immunosuppressive drug *cyclosporin*, a cyclic decapeptide.[143–145] It has a molecular weight of 1.2 kDa, five hydrogen bond donors, 24 acceptors and is highly lipophilic, thus it strongly violates the Pfizer's rule of five. Nevertheless, it fulfils the ADME parameters at best. This is the case for many other known active substances. One explanation that hold for many of these cases is that the drug acts as a substrate for the body's transport systems and that the active substance it is actually a metabolite of the drug.

A new approach to define the properties of a promising drug candidate was formulated by Veber *et al.*.[146] There the rigidity of the molecule and its surface are of prime importance, whereas the size is not a limiting factor. Thus the number of rotateable bonds should be below 10 and the polar surface not exceed 140 Å^2.

8.4 Peptidomimetics

The specificity of $\alpha 4\beta 7$ integrin in LDT sequence recognition made it interesting for more profound investigations on putative inhibitors. As an aromatic residue N-terminal to the LDT sequence proved to be advantageous for biological activity, this position was considered in the investigations. This finally resulted in the highly active lead compound *cyclo*(-Phe-Leu-Asp-Thr-Asp-D-pro-), depicted in Figure 8.3.[147]
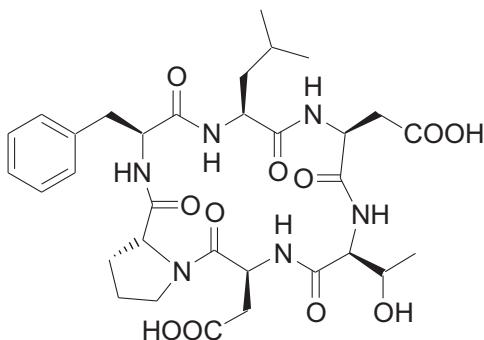
FIGURE 8.3: **The lead-compound *cyclo*(-Phe-Leu-Asp-Thr-Asp-D-pro-).**

Following, several peptidic $\alpha 4\beta 7$ integrin antagonists have been derived from the lead compound *cyclo*(-Phe-Leu-Asp-Thr-Asp-D-pro-).[147, 148] For the further development of non-peptidic and highly selective drugs, the structural and functional requirements of the LDT recognition sequence within $\alpha 4\beta 7$ integrin antagonists have been identified by using peptidomimetic variations like peptoids, azapeptides and reduced amide bonds.

Peptoid-peptide-hybrids show different relative distances between side chains with respect to their position on the backbone. Due to the tertiary amide bonds, the hydrogen binding characteristics of the backbone are changed and the flexibility is increased.[149, 150] The incorporation of azaamino
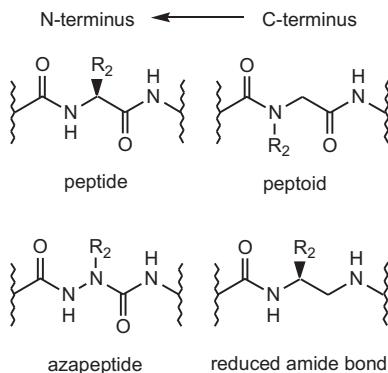
FIGURE 8.4: **Comparison of different peptidomimetics.**

acids maintains the hydrogen bonding characteristics of the backbone. Theoretical studies suggest that azapeptides should be more rigid than their conventional peptide counterparts.[151] As a consequence, azapeptides have the potential to increase activity and selectivity. Furthermore, the incorporation of azaamino acids results in a decrease of the number of stereo centres. The configuration of an azaamino acid lies roughly between the D- and L-configuration of a normal amino acid.[152, 153] The construction of frequently used peptidomimetics is shown in Figure 8.4.

In the lead compound *cyclo*(-Phe-Leu-Asp-Thr-Asp-D-pro-) each amino acid within the FLDT sequence was systematically substituted by the corresponding peptoid building block (Table 8.1). To evaluate the biological influence of the modifications, cell adhesion assays were carried out.[148, 154] The tests showed that not only the correct distances between the functional groups are essential for biological activity, but that the peptide backbone may also influence binding to the receptor. It is conceivable that the amide bonds in the cyclic peptides do not only play an important role in interacting with the integrin, but also, or even exclusively, to stabilise the cyclic

compound	N°	cell adhesion [%]		
		$\alpha 4\beta 7$ / MAdCAM-1	$\alpha 4\beta 7$ / VCAM-1	$\alpha 4\beta 1$ / VCAM-1
c(F-L-D-T-D-p)	1	3 ± 4	38 ± 20	72 ± 10
c(Tic-L-D-T-D-p)	2	9 ± 1	not tested	not tested
c(A-L-D-V-D-p)	3	40 ± 7	not tested	not tested
c(Nphe-L-D-T-D-p)	4	19 ± 10	60 ± 28	97 ± 6
c(F-Nleu-D-T-D-p)	5	102 ± 5	86 ± 13	101 ± 10
c(F-L-Nasp-T-D-p)	6	88 ± 22	94 ± 12	not tested
c(F-L-D-Nval-D-p)	7	98 ± 20	92 ± 19	106 ± 11
c(azaPhe-L-D-T-D-p)	8	25 ± 1	58 ± 7	89 ± 5
c(F-azaLeu-D-T-D-p)	9	87 ± 9	88 ± 9	100 ± 8
c(F-L-azaAsp-T-D-p)	10	106 ± 3	103 ± 12	99 ± 5
c(F-L-D-azaVal-D-p)	11	89 ± 17	not tested	102 ± 8
c(FΨ(CH2NH)-L-D-T-D-p)	12	113 ± 19	102 ± 3	98 ± 9
c(F-LΨ(CH2NH)-D-T-D-p)	13	106 ± 7	100 ± 5	98 ± 0
c(F-L-DΨ(CH2NH)-T-D-p)	14	114 ± 13	108 ± 2	99 ± 7
c(F-L-D-TΨ(CH2NH)-D-p)	15	36 ± 19	95 ± 5	78 ± 19

TABLE 8.1: **Specific effects of peptidomimetics (1 mg/mL) containing a peptoid building block on $\alpha 4\beta 7$- and $\alpha 4\beta 1$-integrin mediated cell adhesion to MAdCAM-1 and VCAM-1.** Cell adhesion is presented as a percentage of medium control. The data represents the mean SD of three experiments, repeated three times. Table taken from Gottschling *et al.*.[148]

structure via hydrogen bonds. The corresponding linear peptides were often inactive whereas the cyclic peptides maintained activity, thus it seems reasonable that the LDT conformation in the cyclic form is very close to the bioactive conformation. Notably, the activity of the constrained mimic *cyclo*(-Tic-Leu-Asp-Thr-Asp-D-pro-) (1) did not decrease significantly in comparison to the lead compound *cyclo*(-Phe-Leu-Asp-Thr-Asp-D-pro-) (2).

8.5 NMR Investigations on cyclo(-Tic^1-Leu^2-Asp^3-Thr^4-Asp^5-D-pro^6-)

Interestingly, compound N° 2 in Table 8.1, *cyclo*(-Tic-Leu-Asp-Thr-Asp-D-pro-), was the only one to show almost as high affinity to MAdCAM-1 as the lead compound N° 1. The structure of the cyclo-hexapeptide *cyclo*(-Leu-Asp-Thr-Ala-D-pro-Phe-), which is similar to lead compound 1 had been determined previously.[147] Though the structure showed D-proline as expected in the $i + 1$ position of a βII'-turn, it was not possible to tell from the data which position the aromatic side chain of the phenylalanine occupies in the active form. It seemed to be highly mobile in solution as only three sequential backbone ROE cross peaks could be observed for phenylalanine.

As the aromatic residue in N-terminal position of the LDT motive had been shown to increase activity considerably, it is important for further in-
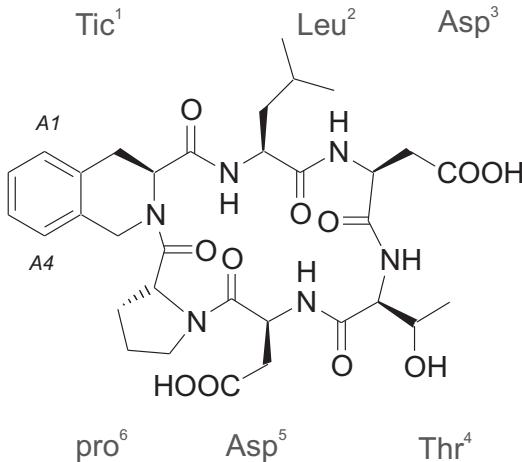
FIGURE 8.5: **Structural formula of *cyclo*(-Tic^1-Leu^2-Asp^3-Thr^4-Asp^5-D-pro^6-).**

vestigations to know about its exact position. Substitution of phenylalanine in the lead compound by the constrained mimic 1,2,3,4-tetrahydro-isoquinoline-3-carboxylic acid (Tic) resulted in an highly active peptide (Figure 8.5). As the flexibility of the aromatic side chain is reduced, the orientation of the aromatic ring, which proved to be important for the biological activity, can be determined. The aromatic side chain must be fixed in a position resembling the one of the bioactive form. This was persuasively demonstrated by the cell adhesion assays as the implementation of this highly constrained phenylalanine mimic did not lead to an significant decrease in binding affinity to MAdCAM-1 (see Table 8.1).

8.5.1 Resonance Assignment

The choice of the appropriate solvent affords careful consideration. In routine NMR spectroscopy of peptides DMSO-d_6 is the favoured solvent, as it combines good solubility with the advantage that no solvent suppression is required. One disadvantage of DMSO are its highly coordinative properties, thus it is know to stabilise βII- and βII'-turns in peptides.[155] In order to avoid any artificial influence on the structure, the NMR experiments were chosen to be carried out in aqueous solution ($H_2O : D_2O = 9 : 1$).

Surprisingly the cyclic hexapeptide was not stable in solution when exposed to air. To avoid both, oxidative decomposition of the peptide and quenching of the ROE signals by solvated oxygen, the sample was shock frozen in liquid nitrogen and then allowed to warm to room temperature while being evacuated with a turbo molecular pump. This procedure was repeated three times to ensure complete evacuation of the sample. Eventually the NMR tube was sealed by melting the tip off. Under strict exclusion of air, the sample proved to be stable over a sufficient long period.

To enable the acquisition of an ^{1}H,^{15}N-HSQC at natural abundance, a relatively high concentration of peptide (22 mM) was chosen. The measurements were performed at 300 K.

	Tic[1]‡	Leu[2]	Asp[3]	Thr[4]	Asp[5]	D-pro[6]
H^N	–	7.38	8.73	7.93	7.58	–
N^H	*n.a.*	122.97	119.69	109.85	119.22	*n.a.*
CO	174.68	175.40	176.11	173.24	171.08	*n.a.*
H^α	5.10	4.36	4.32	4.25	5.19	5.07
C^α	58.08	52.75	55.04	62.39	50.47	59.35
$H^{\beta 1/2}$	3.17/3.36	1.31/1.40*	2.95/3.02*	4.46	2.80/2.96*	2.35/2.04
C^β	33.48	42.61	36.90	68.99	39.06	31.19
$H^{\gamma 1/2}$	–	0.30	–	1.19	–	2.10/2.24*
C^γ	–	25.87	–	21.81	–	27.64
$H^{\delta 1/2}$	–	0.41/0.66*	–	–	–	3.92/3.79
C^δ	–	22.31/24.85*	–	–	–	50.73
$H^{\epsilon 1/2}$	5.30/5.32*	–	–	–	–	–
C^ϵ	49.2	–	–	–	–	–

TABLE 8.2: **Chemical shift assignment of *cyclo*(-Tic[1]-Leu[2]-Asp[3]-Thr[4]-Asp[5]-D-pro[6]-).** * no stereospecific assignment available. *n.a.* chemical shift assignment not available. ‡ The chemical shifts for the aromatic protons A1 and A4 were assigned to 7.28 and 7.37 ppm, respectively. For atom numbering see Figure 8.5.

Proton chemical shifts were assigned with a combination of a TOCSY and a COSY experiment. Proton bound carbon atoms were assigned subsequently with an ^{1}H,^{13}C-HSQC spectrum, quaternary carbon atoms via an ^{1}H,^{13}C-HMBC experiment. Sequential assignment of the two aspartate residues was achieved by evaluating the sequential H^N–H^α ROE cross peaks from the ROESY spectrum. Stereospecific assignment of was equally obtained by evaluation of the ROESY cross peaks. For proline the resonances of $H^{\beta 1}$, $H^{\beta 2}$, $H^{\delta 1}$ and $H^{\delta 2}$ and for the Tic residue $H^{\beta 1}$ and $H^{\beta 2}$ could be explicitly assigned.

8.5.2 ROESY Spectra for Distance Restraints

Due to the low molecular weight of only 648 Da of the cyclic hexapeptide, a ROESY experiment was favoured instead of a NOESY experiment.

The ROESY spectrum was acquired at three different mixing times, $\tau_{mix} =$ 150, 200, 250 ms. The spectrum with a mixing time of 150 milliseconds showed a insufficient ROE build-up. The spectra with mixing times of 200 and 250 milliseconds did not show a significant difference in cross peak intensity. In order to minimize ambiguous information from cross peaks due to spin diffusion the spectrum with $\tau_{mix} = 200$ ms was chosen for further evaluation (Figure 8.6). The integration of the cross peaks was done with the X-WINNMR software. If possible both, the peaks above and below the diagonal were evaluated and the average was taken, though this was hindered for peaks in proximity of the water resonance at 4.7 ppm due to the water suppression scheme.

Diastereotopic protons, e.g. $H^{\beta 1}$ and $H^{\beta 2}$, are ideal for the calibration of the integrals as their distance is almost independent of the dihedral angle. Unfortunately all possible $H^{\beta 1/2}$ from Leu^2, Asp^3 and Asp^5 show quite similar chemical shifts and thus their cross peaks were buried or heavily overlapped with noise originating from the diagonal making a quantitative integration impossible. For the calibration of the integrals the distances from H^{α} to $H^{\beta 1}$ and $H^{\beta 2}$ from proline were chosen. They were set to 2.36 and 3.05 Å, respectively.

In total 86 contacts were evaluated of which 40 were biased as they overlapped either with other resonances, with the water signal or the diagonal. Of the remaining 46 contacts 27 were not redundant and thus were included into the final calculation of the structure. 10 of the distance restraints arouse from inter-residue contacts, 5 were sequential H^N–H^{α} contacts and 12 were caused by medium to long range contacts. A list of the restraints used for the calculations is given in Table 8.3. As a standard, a deviation of $\pm 10\%$ from the calculated value was allowed.
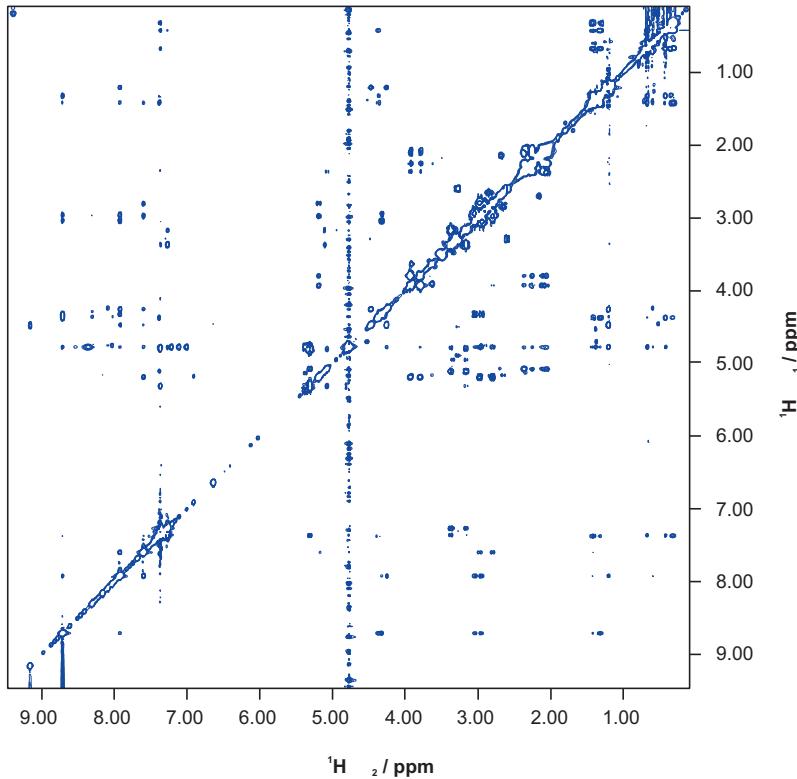
FIGURE 8.6: **ROESY spectrum of** *cyclo***(-Tic¹-Leu²-Asp³-Thr⁴-Asp⁵-D-pro⁶-).** The spectrum was acquired at 300 K on a 500 MHz spectrometer with an ROE mixing time of $\tau_{mix} = 200$ ms and a pulsed spin-lock with a field strength of 3180 Hz.

	Distance		lower bound	upper bound	calc. Distance	Viol.
1	Tic^1 A1	Tic^1 H$^{\beta 1}$	2.18	2.66	2.48	0
2	Tic^1 A1	Tic^1 H$^{\beta 2}$	2.80	3.42	3.19	0
3	Tic^1 H$^{\beta 1}$	Tic^1 H$^\epsilon$	3.56	5.17	3.98*	0
4	Tic^1 H$^{\beta 2}$	Tic^1 H$^\epsilon$	2.67	4.08	3.23*	0
5	Leu^2 H^N	Leu^2 H$^\alpha$	2.60	3.18	3.03	0
6	Leu^2 H^N	Asp^5 H^N	4.07	4.97	4.77	0
7	Leu^2 H^N	Tic^1 H$^\alpha$	3.02	3.69	3.46	0
8	Leu^2 H^N	Tic^1 H$^{\beta 1}$	3.92	4.79	3.88	0.04
9	Asp^3 H^N	Leu^2 H$^\alpha$	2.02	2.46	2.14	0
10	Asp^3 H^N	Leu^2 H$^\beta$	2.77	4.91	3.62*	0
11	Asp^3 H^N	Asp^3 H$^\alpha$	2.11	2.58	2.31	0
12	Asp^3 H^N	Thr^4 H^N	3.06	3.73	3.46	0
13	Asp^3 H^N	Asp^5 H^N	4.65	5.68	5.20	0
14	Thr^4 H^N	Leu^2 H$^\beta$	3.37	5.60	5.61*	0.01
15	Thr^4 H^N	Asp^3 H$^\alpha$	2.51	3.07	2.66	0
16	Thr^4 H^N	Asp^3 H$^\beta$	2.43	4.44	3.33*	0
17	Thr^4 H^N	Thr^4 H$^\alpha$	2.31	2.82	2.50	0
18	Thr^4 H^N	Thr^4 H$^\beta$	2.85	3.48	3.11	0
19	Thr^4 H^N	Thr^4 H$^{\gamma 2}$	2.46	4.53	3.64*	0
20	Thr^4 H^N	Asp^5 H^N	2.35	2.88	2.86	0
21	Asp^5 H^N	Leu^2 H$^\beta$	3.62	5.95	5.85*	0
22	Asp^5 H^N	Asp^3 H$^\alpha$	4.38	5.36	4.85	0
23	Asp^5 H^N	Thr^4 H$^\alpha$	3.16	3.87	3.40	0
24	Asp^5 H^N	Thr^4 H$^\beta$	3.61	4.42	4.19	0
25	Asp^5 H^N	Asp^5 H$^\alpha$	2.58	3.15	3.04	0
26	Asp^5 H$^\alpha$	D-pro^6 H$^\delta$	2.11	3.28	2.62*	0
27	D-pro^6 H$^\alpha$	Tic^1 H$^\epsilon$	1.81	3.73	2.59*	0

TABLE 8.3: **Distance restraints derived from the ROESY spectrum that were used for the final structure calculation for** *cyclo***(-Tic^1-Leu^2-Asp^3-Thr^4-Asp^5-D-pro^6-).** Violations correspond to those which occurred during the 300 picosecond time averaged distance restraints MD simulation and were calculated by $< r^{-3} >^{-1/3}$ averaging. * Not stereospecifically assigned. For the calculation, the carbon atom to which the proton is bonded to, was used.

8.5.3 Temperature Coefficients

Measurement of the temperature dependence of the ^{1}H chemical shifts of amide protons in peptides is a well established method to estimate how exposed the amide proton is to the solvent.[156] The temperature coefficients of *cyclo*(-Tic-Leu-Asp-Thr-Asp-D-pro-) were determined in steps of 5 degrees between 285 and 310 K. The temperature dependence of the four amide protons of *cyclo*(-Tic-Leu-Asp-Thr-Asp-D-pro-)are shown in Figure 8.7, the according temperature coefficients of are given below in Table 8.4.
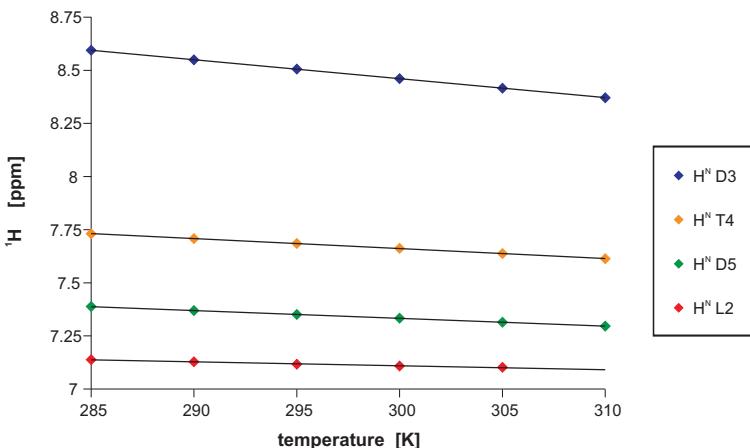
FIGURE 8.7: **Temperature dependence of the amide protons for *cyclo*(-Tic^1-Leu^2-Asp^3-Thr^4-Asp^5-D-pro^6-).**

residue	Leu^2	Asp^3	Thr^4	Asp^5
$\Delta\delta$ (H^N) / ΔT	−1.8	−8.9	−4.7	−3.7

TABLE 8.4: **Temperature dependence of the amide protons of *cyclo*(-Tic^1-Leu^2-Asp^3-Thr^4-Asp^5-D-pro^6-).** Coefficients are given in ppb/K.

Asp[3] shows the highest coefficient with -8.9 ppb/K. The coefficients for Thr[4] and Asp[5] are in a medium range -4.7 and -3.7 ppb/K, respectively. Contrarily, Leu[2] is almost unaffected with -1.8 ppb/K. This suggest that Leu[2] is strongly hydrogen bonded. Weaker hydrogen bonding can be predicted for Asp[5] and Thr[4], whereas the amide proton of Asp[3] seems to be in rapid exchange with the water protons according to its high temperature dependency and thus is with high probability completely exposed to the solvent.

8.5.4 The Solution Structure of cyclo(-Tic[1]-Leu[2]-Asp[3]-Thr[4]-Asp[5]-D-pro[6]-)

Distance geometry (DG) calculations[157] with the distance restraints obtained from the ROESY spectrum and molecular dynamics (MD) simulations[158] show that D-pro[6] occupies position $i + 1$ in a βII'-turn. The data suggest a flip of the peptide bond between Asp[3] and Thr[4], indicating that there might be at least two conformations. Using a time averaged distance restraints protocol,[159] the flip of the peptide bond between Asp[3] and Thr[4] could clearly be observed after approximately 130 picoseconds of simulation time. Averaging of the two conformations, and subsequent minimization yielded the structures depicted in Figure 8.8. The position of the aromatic ring of Tic[1], which is very important for the biological activity, is nearly perpendicular to the peptide backbone. The spatial position of Tic[1] is equal in both conformers.

The calculated structure is in accordance with the previously obtained data from the temperature coefficients. Leu[2] occupies position $i + 2$ in the βII'-turn, thus its amide proton is involved in a hydrogen bond to the carbonyl oxygen atom from D-pro[6] in position $i + 1$. The amide protons of Thr[4] and Asp[5] showed a decreased temperature coefficient. Most likely the reason therefore is the flipping of the peptide bond between Asp[3] and Thr[4] that disrupts the hydrogen bonding network. The amide proton of Asp[3]
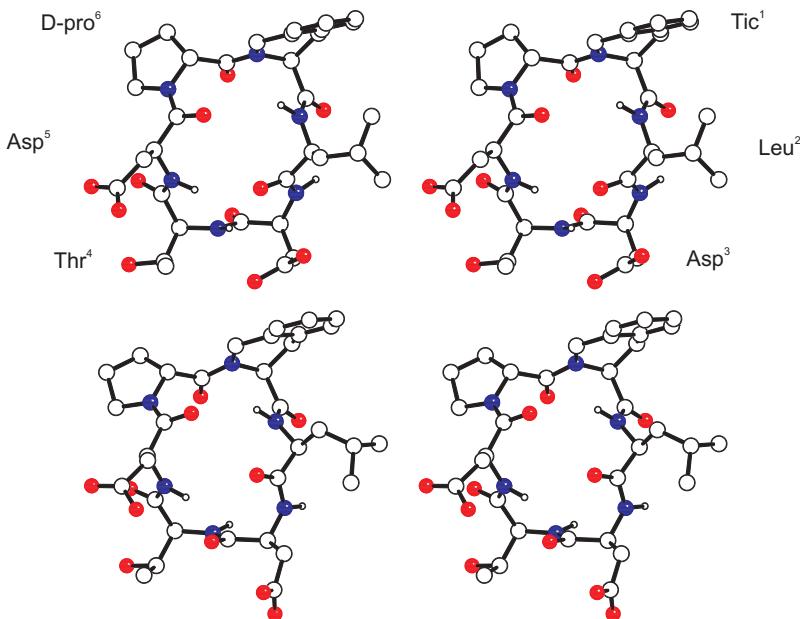
FIGURE 8.8: **Stereo view of the solution structures of the two conformers of *cyclo*(-Tic[1]-Leu[2]-Asp[3]-Thr[4]-Asp[5]-D-pro[6]-).** D-pro[6] occupies position $i + 1$ in a βII'-turn. The spatial position of Tic[1] is equal in both conformers, whereas the peptide bond between Asp[3] and Thr[4] is flipped.

points in both conformations observed into the solvent, making a stabilisation of the proton through hydrogen bond impossible. This explains its high temperature dependence.

These results indicate that the C-terminal aromatic system, which is important for the biological activity, should be perpendicular to the peptide backbone. Furthermore the orientation of the peptide bond between Asp[3] and Thr[4] seems not to be of biological relevance.

In combination with the biological data, the NMR structure provides important information for the design of novel and highly active $\alpha 4\beta 7$ integrin antagonists. Further efforts to reduce the peptidic character of these peptidomimetics can lead to drug candidates for the treatment of inflammatory diseases.

8.6 Materials and Methods

8.6.1 NMR Spectroscopy

Prior to spectral analysis, the prepared sample was shock frozen in the NMR-tube in liquid nitrogen. It was allowed to warm up to room temperature while being evacuated with a turbo molecular pump. Atmospheric pressure was restored and balanced with a nitrogen filled balloon. Carefully, the tube was dipped into liquid nitrogen again. This procedure was repeated three times before the NMR tube was sealed by melting the tip off.

All spectra were acquired on Bruker DMX500 or DMX600 spectrometers in H_2O/D_2O (9:1) as solvent. The measurements were performed at 300 K, using a 22 mM solution of peptide. For the determination of the amide proton and nitrogen shifts an $^1H,^{15}N$-HSQC-spectrum was recorded. Proton chemical shifts were assigned with a COSY and a TOCSY experiment. Carbon resonance assignment was achieved using a $^1H,^{13}C$-HSQC-spectrum in combination with an $^1H,^{13}C$-HMBC experiment. Stereospecific assignment was done, if applicable, via a thorough investigation of the ROE cross peaks. Water suppression was achieved using WATERGATE.[160] The spectra were processed and analysed on Silicon Graphics workstations using the XWIN-NMR and AURELIA software packages.

8.6.2 Structure Determination

Distance information for the structure investigation was derived from a ROESY experiment with a pulsed spin-lock with a field strength of 3180 Hz on a 500 MHz spectrometer and with a mixing time of $\tau_{mix} = 200$ ms. On the basis of complete absence of $H^\alpha(i) - H^\alpha(i+1)$ cross peaks in the ROESY-spectrum, all peptide bonds were assumed to be in *trans*-configuration. Temperature coefficients were determined between 285 and 310 K in steps of 5 degrees and referenced internally to H^γ of Leu^2. The proton-proton distances were calculated by integration of the ROESY cross peaks and a tol-

erance of $\pm 10\%$ was applied to derive lower and upper bounds for the structure calculations. The regions above and below the diagonal of the ROESY spectrum were analysed and compared. 27 interproton distances were taken for the final structure calculation. Distance restraint data and violations are shown in Table 8.3.

8.6.3 Molecular Dynamics

The structure calculations were performed on a Silicon Graphics Origin 200. An initial conformational search was performed with distance geometry (DG) calculations, using a modified version of the DISGEO program package. The structural refinement calculations were done by Luciana Marinelli. An ensemble of 100 distance geometry structures with good convergence was calculated. The best structure was placed in a cubical box with 1935 water molecules and a size of 39 x 39 x 39 Å^3. A dynamics simulation of 300 picoseconds duration, applying the CVFF forcefield, was performed with the DISCOVER program in conjunction with a homemade extension for time averaged distance restraints. The force constant on distance restraints was set to $10 \, \text{kJ} \, \text{mol}^{-1} \, \text{Å}^{-2}$, and a decay constant t of 30 picoseconds was used. The calculation was done with the explicit-image model of periodic boundary conditions. A flip of the peptide bond between Asp^3 and Thr^4 was observed after 133 picoseconds simulation time. The trajectory was divided into two parts, ranging from 1–118 and 149–300 picoseconds, and omitting the time of transition between the two conformations. Each trajectory part was then averaged and minimized with 300 steps of the steepest decent algorithm, yielding the final structures depicted in Figure 8.8.

Bibliography

[1] M. H. Saier, Jr. Vectorial metabolism and the evolution of transport systems. *J. Bacteriol.*, 182:5029–5035, 2000.

[2] L. Stryer. *Biochemie*. Spektrum Akademischer Verlag, Heidelberg, Berlin, Oxford, 4. auflage edition, 1996.

[3] B. M. Denker, B. L. Smith, F. P. Kuhajda, and P. Agre. Identification, purification, and partial charcterization of a novel Mr 28,000 integral membrane protein from erythrocytes and renal tubules. *J. Biol. Chem.*, 263:15634–15642, 1988.

[4] K. Murata, K. Mitsuoka, T. Hirai, T. Walz, P. Agre, J. B. Heymann, A. Engel, and Y. Fujiyoshi. Structural determinants of water permeation through aquaporin-1. *Nature*, 407:599–605, 2000.

[5] H. Sui, B. G. Han, J. K. Lee, P. Walian, and B. K. Jap. Structural basis of water-specific transport through the AQP1 water channel. *Nature*, 414:872–878, 2001.

[6] Q. Ye, B. Wiera, and E. Steudle. A cohesion/tension mechanism explains the gating of water channels (aquaporins) in *Chara internodes* by high concentration. *J. Exp. Bot.*, 55:449–462, 2004.

[7] C. Tournaire-Roux, M. Sutka, H. Javot, E. Gout, P. Gerbeau, D. T. Luu, R. Bligny, and C. Maurel. Cytosolic pH regulates root water transport

during anoxic stress through gating of aquaporins. *Nature*, 425:393–397, 2003.

[8] X. Wan, E. Steudle, and W. Hartung. Gating of water channels (aquaporins) in cortical cells of young corn roots by mechanical stimuli (pressure pulses): effects of ABA and of $HgCl_2$. *J. Exp. Bot.*, 55:411–422, 2004.

[9] W. Ostwald. Elektrische Eigenschaften halbdurchlässiger Scheidewände. *Z. Phys. Chem.*, 6:71–82, 1890.

[10] L. Michaelis. Contribution to the theory of permeability of membranes for electrolytes. *J. Gen. Physiol.*, 8:33–59, 1925.

[11] A. L. Hodgkin and R. D. Keynes. The potassium permeability of a giant nerve fibre. *J. Physiol.*, 128:61–88, 1955.

[12] J. Kistler, R. M. Stroud, M. W. Klymkowsky, R. A. Lalancette, and R. H. Fairclough. Structure and function of an acetylcholine receptor. *Biophys. J.*, 37:371–383, 1982.

[13] M. J. Ross, M. W. Klymkowsky, D. A. Agard, and R. M. Stroud. Structural studies of a membrane-bound acetylcholine receptor from *Torpedo californica*. *J. Mol. Biol.*, 116:635–659, 1977.

[14] M. Noda, S. Shimizu, T. Tanabe, T. Takai, T. Kayano, T. Ikeda, H. Takahashi, H. Nakayama, Y. Kanaoka, and N. M. et al. Primary structure of *Electrophorus electricus* sodium channel deduced from cDNA sequence. *Nature*, 312:121–127, 1984.

[15] Tetrodoxin Is Safe and Effective for Severe, Refractory Cancer Pain. *J. Supportive Oncology*, 18:2, 2004.

[16] D. A. Doyle, J. M. Cabral, R. A. Pfuetzner, A. Kuo, J. M. Gulbis, S. L. Cohen, B. T. Chait, and R. MacKinnon. The structure of the potassium

channel: molecular basis of K$^+$ conduction and selectivity. *Science*, 280:69–77, 1998.

[17] J. H. Morais-Cabral, Y. Zhou, and R. MacKinnon. Energetic optimizytion of ion conduction rate by the K$^+$ selectivity filter. *Nature*, 414:37–42, 2001.

[18] J. C. Skou. The influence of some cations on an adenosine triphosphatase from peripheral nerves. *Biochim. Biophys. Acta*, 23:394–401, 1957.

[19] J. V. Møller, B. Juul, and M. le Maire. Structural organization, ion transport, and energy transduction of P-type ATPases. *Biochim. Biophys. Acta*, 1286:1–51, 1996.

[20] P. L. Pedersen and E. Carafoli. Ion motive ATPases. I. Ubiquity, properties, and significance to cell function. *Trends Biochem. Sci.*, 12:146–150, 1987.

[21] K. B. Axelsen and M. G. Palmgren. Evolution of substrate specificities in the P-type ATPase superfamily. *J. Mol. Evol.*, 46:84–101, 1998.

[22] R. Serrano. Structure and function of proton translocation ATPase in plasma membranes of plants and fungi. *Biochim. Biophys. Acta*, 947:1–28, 1988.

[23] C. Toyoshima, M. Nakasako, H. Nomura, and H. Ogawa. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature*, 405:647–655, 2000.

[24] K. Altendorf and W. Epstein. *The Kdp-ATPase of* Escherichia coli, volume 5, pages 403–420. JAI Press Inc., Greenwich London, 1996.

[25] M. Bramkamp. *Characterization of the KdpFABC complex from* Escherichia coli, *of soluble subdomains from KdpB, and of a homologous protein of* Methanococcus jannaschii. PhD thesis, Universität Osnabrück, 2003.

[26] E. T. Buurman, K.-T. Kim, and W. Epstein. Genetic evidence of two sequentially occupied K^+ binding sites in the Kdp transport ATPase. *J. Biol. Chem.*, 270:6678–6685, 1995.

[27] S. R. Durell, E. P. Bakker, and H. R. Guy. Does the KdpA subunit from the high affinity K^+-translocationg P-type Kdp-ATPase have a structure similar to that of K^+ channels? *Biophys. J.*, 78:188–199, 2000.

[28] A. Siebers and K. Altendorf. *Alkali cation transport systems in prokaryotes*, chapter K^+-translocating Kdp-ATPases and other bacterial P-type ATPases, pages 225–252. CRC Press, Boca Raton Florida, 1993.

[29] W. Puppe, A. Siebers, and K. Altendorf. The phosphorylation site of the Kdp-ATPase of *Escherichia coli*: site-directed mutageniesis of the aspartic acid residues 300 and 307 of the KdpB subunit. *Mol. Microbiol.*, 6:3511–3520, 1992.

[30] K. Altendorf, A. Siebers, and W. Epstein. The Kdp-ATPase of *Escherichia coli*. *Ann. N Y Acad. Sci.*, 671:228–243, 1992.

[31] K. A. Taylor, L. Dux, and A. Martonosi. Three-dimensional reconstruction of negatively stained crystals of the Ca^{2+}-ATPase from muscle sarcoplasmic reticulum. *J. Mol. Biol.*, 187:417–427, 1986.

[32] A. Siebers and K. Altendorf. Characterization of the phosphorylated intermediate of the K^+-translocating Kdp-ATPase form *Escherichia coli*. *J. Biol. Chem.*, 264:5831–5838, 1989.

[33] J. D. Clausen, B. Vilsen, D. B. McIntosh, A. P. Einholm, and J. P. Andersen. Glutamate-183 in the conserved TGES motif of domain A of sarcoplasmic reticulum Ca^{2+}-ATPase assists in catalysis of E2/E2P partial reactions. *Proc. Natl. Acad. Sci. USA*, 101:2776–2781, 2004.

[34] M. Haupt, M. Coles, M. Bramkamp, K. Altendorf, and H. Kessler. Inter-domain Motions of the N-domain of the KdpFABC Complex,

a P-type ATPase, are not Driven by ATP-induced Conformational Changes. *J. Mol. Biol.*, 342:1547–1558, 2004.

[35] M. Gaßel and K. Altendorf. Analysis of KdpC of the K$^+$-transporting KdpFABC complex of *Escherichia coli*. *Eur. J. Biochem.*, 268:1772–1781, 2001.

[36] K. Altendorf, M. Gaßel, W. Puppe, T. Möllenkamp, A. Zeeck, C. Boddien, K. Fendler, E. Bamberg, and S. Dröse. Structure and function of the Kdp-ATPase of *Escherichia coli*. *Acta Physiol. Scand. Suppl.*, 163:137–146, 1998.

[37] M Gaßel and T Möllenkamp and W Puppe and K Altendorf. The KdpF subunit is part of the K$^+$-translocating Kdp complex of *Escherichia coli* and is responsible for the stabilization of the complex *in vitro*. *J. Biol. Chem.*, 274:37901–37907, 1999.

[38] P. Zimmann, W. Puppe, and K. Altendorf. Membrane Topology Analysis of the Sensor Kinase KdpD of *Escherichia coli*. *J. Biol. Chem.*, 270:28282–28288, 1995.

[39] P. Voelkner, W. Puppe, and K. Altendorf. Characterization of the KdpD protein, the sensor kinase of the K$^+$-translocating Kdp system of *Escherichia coli*. *Eur. J. Biochem.*, 217:1019–1026, 1993.

[40] W. Puppe, P. Zimmann, K. Jung, M. Lucassen, and K. Altendorf. Characterization of Truncated Forms of the KdpD Protein, the Sensor Kinase of the K$^+$-translocating Kdp System of *Escherichia coli*. *J. Biol. Chem.*, 271:25027–25034, 1996.

[41] K. Jung and K. Altendorf. Individual Substitutions of Clustered Arginine Residues of the Sensor Kinase KdpD of *Escherichia coli* Modulate the Ratio of Kinase to Phosphatase Activity. *J. Biol. Chem.*, 273:26415–26420, 1998.

[42] M. O. Walderhaug, J. W. Polarek, P. Voelkner, J. M. Daniel, J. E. Hesse, K. Altendorf, and W. Epstein. KdpD and KdpE, proteins that control expression of the *kdpABC* operon, are members of the two-component sensor-effector class of regulators. *J. Bacteriol.*, 174:2152–2159, 1992.

[43] M. Esmann and J. C. Skou. The effect of K^+ on the equilibrium between the E2 and the K-occluded E2 conformation of the Na^+,K^+-ATPase. *Biochim. Biophys. Acta*, 748:413–417, 1983.

[44] R. W. Albers. Biochemical aspects of active transport. *Annu. Rev. Biochem.*, 36:727–756, 1967.

[45] R. L. Post, C. Hegyvary, and S. Kume. Activation by adenosine triphosphate in the phosphorylation kinetics of sodium and potassium ion transport adenosine triphosphatase. *J. Biol. Chem.*, 247:6530–6540, 2001.

[46] L. DeMeis and A. L. Vianna. Energy interconversion by the Ca^{2+}-dependent ATPase of the sarcoplasmic reticulum. *Ann. Rev. Biochem.*, 48:275–292, 2002.

[47] L. Pauling. Molecular architecture and biological reactions. *Chem. Eng. News*, 24:1375–1377, 1946.

[48] G. A. Scarborough. Molecular Mechanism of the P-type ATPases. *J. Bioenerg. Biomembr.*, 34:235–250, 2002.

[49] P. Zhang, C. Toyoshima, K. Yonekura, N. M. Green, and D. Stokes. Structure of the calcium pump from sarcoplasmic reticulum at 8 Å resolution. *Nature*, 392:835–839, 1998.

[50] M. Auer, G. A. Scarborough, and W. Kühlbrand. Three-dimensional map of the plasma membrane H^+-ATPase in the open conformation. *Nature*, 392:840–843, 1998.

[51] P. L. Jørgensen. Purification and characterization of Na$^+$,K$^+$-ATPase. V. Conformational changes in the enzyme Transitions between the Na-form and the K-form studied with tryptic digestion as a tool. *Biochim. Biophys. Acta*, 401:399–415, 1975.

[52] A. Rephaeli, D. Richards, and S. J. Karlish. Conformational transitions in fluorescein-labeled (Na,K)ATPase reconstituted into phospholipid vesicles. *J. Biol. Chem.*, 261:6248–6254, 1986.

[53] C. Toyoshima and H. Nomura. Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature*, 418:605–611, 2002.

[54] W. J. Rice, H. S. Young, D. W. Martin, J. R. Sachs, and D. L. Stokes. Structure of the Na$^+$,K$^+$-ATPase at 11 Å Resolution: Comparison with Ca^{2+}-ATPase in E1 and E2 States. *Biophys. J.*, 80:2187–2197, 2001.

[55] H. Hebert, P. P. an H Vorum an K Thomsen, and A. B. Maunsbach. Three-dimensional structure of renal Na$^+$,K$^+$-ATPase from cryo-electron microscopy of two-dimensional crystals. *J. Mol. Biol.*, 314:479–494, 2001.

[56] W. Kühlbrandt, J. Zeelen, and J. Dietrich. Structure, Mechanism, and Regulation of the *Neurospora* plasma membrane H$^+$-ATPase. *Science*, 297:1692–1696, 2002.

[57] A. G. Lee. Ca^{2+}-ATPase structure in the E1 and E2 conformations: mechanism, helix-helix and helix-lipid interactions. *Biochim. Biophys. Acta*, 1565:246–266, 2002.

[58] A. G. Lee. A calcium pump made visible. *Curr. Opin. Struct. Biol.*, 12:548–554, 2002.

[59] C. R. D. Lancaster. A P-type ion pump at work. *Nat. Struct. Biol.*, 9:643–645, 2002.

[60] C. Toyoshima, H. Nomura, and Y. Sugita. Structural basis of ion pumping by Ca^{2+}-ATPase of sarcoplasmic reticulum. *FEBS Letters*, 555:106–110, 2003.

[61] M. Bramkamp and K. Altendorf. Mutational analysis of charged residues in the putative KdpB-TM5 domain of the Kdp-ATPase of *Escherichia coli*. *Ann. N. Y. Acad. Sci.*, 986:351–353, 2003.

[62] F. Bloch, W. W. Hansen, and M. Packard. Nuclear Induction. *Phys. Rev.*, 69:127, 1946.

[63] E. M. Purcell, H. C. Torrey, and R. V. Pound. Resonance absorption by nuclear magnetic moments in a solid. *Phys. Rev.*, 69:37–38, 1946.

[64] R. R. Ernst and W. A. Anderson. Application of Fourier transform spectroscopy to magnetic resonance. *Rev. Sci. Instr.*, 37:93–102, 1966.

[65] K. Wüthrich. The way to NMR structures of proteins. *Nat. Struct. Biol.*, 8:923–925, 2001.

[66] M. Hesse, H. Meier, and B. Zeeh. *Spektroskopische Methoden in der organischen Chemie*. Georg Thieme Verlag, Stuttgart, 1995.

[67] N. A. Farrow, R. Muhandiram, A. U. Singer, S. M. Pascal, C. M. Kay, G. Gish, S. E. Shoelson, T. Pawson, J. D. Forman-Kay, and L. E. Kay. Backbone Dynamics of a Free and a Phosphopeptide-Complexed Src Homology-2 Domain Studied by ^{15}N NMR Relaxation. *Biochemistry*, 33:5984–6003, 1994.

[68] V. F. Bystrov. Coupling constants in proteins. *Prog. Nucl. Magn. Reson. Spectrosc.*, 10:41–81, 1976.

[69] M. Sattler, J. Schleucher, and C. Griesinger. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.*, 34:93–158, 1999.

[70] J. N. S. Evans. *Biomolecular NMR Spectroscopy*. Oxford University Press, Oxford, New York, Tokyo, 1996.

[71] M. H. Levitt. *Spin Dynamics*. John Wiley & Sons, LTD, Chichester, New York, Weinheim, Brisbane, Singapore, Toronto, 2002.

[72] O. Zhang and J. D. Forman-Kay. NMR studies of unfolded states of an SH3 domain in aqueous solution and denaturing conditions. *Biochemistry*, 36:3959–3970, 1997.

[73] T. Diercks, M. Coles, and H. Kessler. An efficient strategy for assignment of cross-peaks in 3D heteronuclear NOESY experiments. *J. Biomol. NMR.*, 15:177–180, 1999.

[74] G. Vriend. WHAT IF: A molecular modeling and drug design program. *J. Mol. Graph.*, 8:52–56, 1990.

[75] T. A. Jones and M. Kjeldgaard. *Essential 'O', software manual*. University, Uppsala, Sweden, 1998.

[76] M. Tönnis. *Untersuchungen zur molekularen Interaktion und Zellkommunikation am Beispiel der Inkompatibilitätsloki von* Ustilago myadis. PhD thesis, Universität München, 2002.

[77] J. Sambrock, E. F. Fritsch, and T. Maniatis. *Molecular Cloning. A Laboratory Manual*. Cold Spring Laboratory Press, Cold Spring Harbor, New York, 1989.

[78] M. Leutner, R. M. Gschwind, J. Liermann, C. Schwarz, G. Gemmecker, and H. Kessler. Automated backbone assignment of labeled proteins using the threshold accepting algorithm. *J. Biomol. NMR*, 11:31–43, 1998.

[79] M. Haupt, M. Coles, V. Truffault, M. Bramkamp, K. Altendorf, and H. Kessler. Letter to the Editor: ^{1}H, ^{13}C and ^{15}N Resonance Assignment of the Nucleotide Binding Domain of KdpB from *Escherichia Coli*. *J. Biomol. NMR*, 29:437–438, 2004.

[80] D. S. Wishart and B. D. Sykes. Chemical-Shifts as a Tool for Structure Determination. *Methods Enzymol.*, 239:363–392, 1994.

[81] D. S. Wishart, B. D. Sykes, and F. M. Richards. The Chemical-Shift-Index – A fast and simple method for the assignment of protein secondary structure through NMR-spectroscopy. *Biochemistry*, 31:1647–1651, 1992.

[82] D. S. Wishart and B. D. Sykes. The C-13 Chemical-Shift-Index – A simple method for the identification of protein secondary structure using C-13 chemical shift data. *J. Biomol. NMR*, 4:171–180, 1994.

[83] D. S. Wishart, C. G. Bigam, J. Yao, F. Abildgaard, J. H. Dyson, E. Old-field, J. L. Markley, and B. D. Sykes. ^{1}H, ^{13}C and ^{15}N Chemical-Shift Referencing in Biomolecular NMR. *J. Biomol. NMR*, 6:135–140, 1995.

[84] G. W. Vuister and A. Bax. Quantitative J correlation: a new approach for measuring homonuclear three-bond J(H^N,H$^\alpha$) coupling constants in ^{15}N-enriched proteins. *J. Am. Chem. Soc.*, 115:7772–7777, 1993.

[85] R. A. Laskowski, M. W. MacArthur, D. S. Moss, and J. M. Thornton. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.*, 26:283–289, 1993.

[86] Y. Chen, J. Reizer, M. H. Saier, Jr, W. J. Fairbrother, and P. E. Wright. Mapping of the Binding Interfaces of the Proteins of the Bacterial Phosphotransferase System, HPr and IIA^{glc}. *Biochemistry*, 32:32–37, 1993.

[87] D. S. Garrett, Y.-J. Seok, A. Peterkofsky, G. M. Clore, and A. M. Gronenborn. Identification by NMR of the Binding Surface for the Histidine-Containing Phosphocarrier Protein HPr on the N-Terminal Domain of Enyzme I of the *Escherichia coli* Phosphotransferase System. *Biochemistry*, 36:4393–4398, 1997.

[88] T. Tanaka, K. S. Soumitra, C. Tomomori, R. Ishima, D. Liu, K. I. Tong, H. Park, R. Dutta, L. Qin, M. B. Swindells, R. Yamazaki, A. M. Ono, M. Kainosho, M. Inouye, and M. Ikura. NMR structure of the histidine kinase domain of the *E. coli* osmosensor EnvZ. *Nature*, 396:88–92, 1998.

[89] E. G. Moczydlowski and P. A. Fortes. Characterization of 2',3'-O-(2,4,6-trinitrocyclohexadienylidene) adenosine 5'-triphosphate as a fluorescent probe of the ATP site of sodium and potassium transport adenosine triphosphatase. Determination of nucleotide binding stoichiometry and ion-induced changes in affinity for ATP. *J. Biol. Chem.*, 256:2346–1256, 1981.

[90] E. G. Moczydlowski and P. A. Fortes. Inhibition of the sodium and potassium adenosine triphosphatase by 2',3'-O-(2,4,6-trinitrocyclohexadienylidene) adenine nucleotides. Implications for the structure and mechanism of the Na:K pump. *J. Biol. Chem.*, 256:2357–2366, 1981.

[91] Y. Dupont, Y. Chapron, and R. Pougeois. Titration of the nucleotide binding sites of sarcoplasmic reticulum Ca^{2+}-ATPase with 2',3'-O-(2,4,6-trinitrophenyl) adenosine 5'-triphosphate and 5'-diphosphate. *Biochem. Biophys. Res. Commun.*, 106:1272–1279, 1982.

[92] M. Pellecchia, P. Sebbel, U. Hermanns, K. Wüthrich, and R. Glockshuber. Pilus chaperone FimC-adhesin FimH interactions mapped by TROSY-NMR. *Nat. Struct. Biol.*, 6:336–339, 1999.

[93] M. Abu-Abed, T. K. Mal, M. Kainosho, D. H. MacLennan, and M. Ikura. Characterization of the ATP-Binding Domain of the Sarco-(endo)plasmic Reticulum Ca^{2+}-ATPase: Probing Nucleotide Binding by Multidimensional NMR. *Biochemistry*, 41:1156–1164, 2002.

[94] V. Truffault, M. Coles, T. Diercks, K. Abelmann, S. Eberhardt, H. Lüttgen, A. Bacher, and H. Kessler. The Solution Structure of the N-terminal Domain of Riboflavin Synthase. *J. Mol. Biol.*, 309:949–960, 2001.

[95] V. Truffault. *Solution State NMR Investigations on the Structure and Function of Proteins: Application to Riboflavin Synthase*. PhD thesis, Universität München, 2002.

[96] M. Mayer and B. Meyer. Characterization of Ligand Binding by Saturation Transfer Difference NMR Spectroscopy. *Angew. Chem. Int. Ed.*, 38:1784–1788, 1999.

[97] M. Mayer and B. Meyer. Group Epitope Mapping by Saturation Transfer Difference NMR to Identify Segments of a Ligand in Direct Contact with a Protein Receptor. *J. Am. Chem. Soc.*, 123:6108–6117, 2001.

[98] W. Epstein and B. S. Kim. Potassium transport loci in Escherichia coli K-12. *J. Bacteriol.*, 108:639–644, 1971.

[99] A. T. Brünger. *X-PLOR, Version 3.1. A System for X-ray Crystallography and NMR*. Yale University Press, New Haven, Connecticut, 1992.

[100] G. Cornilescu, F. Delaglio, and A. Bax. Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR*, 13:289–302, 1999.

[101] W. L. DeLano. *The PyMOL Molecular Graphics System. Version 0.95*. DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org, 2004.

[102] P. J. Kraulis. MOLSCRIPT - a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallog.*, 24:946–950, 1991.

[103] E. A. Merritt and D. J. Bacon. Raster 3D: Photorealistic molecular graphics. *Meth. Enzymol.*, 277:505–524, 1997.

[104] I. MDL Information Systems. *MDL ISIS^{TM}/Draw 2.5*. MDL Information Systems, Inc., http://www.mdli.com, 2002.

[105] M. Hilge, G. Siegal, G. W. Vuister, P. Guntert, S. M. Gloor, and J. P. Abrahams. ATP-induced conformational changes of the nucleotide-binding domain of Na$^+$,K$^+$-ATPase. *Nat. Struct. Biol.*, 10:468–474, 2003.

[106] K. O. Håkansson. The crystallographic structure of Na,K-ATPase N-domain at 2.6 Å resolution. *J. Mol. Biol.*, 332:1175–1182, 2003.

[107] L. Holm and C. Sander. Protein Structure Comparison by Alignment of Distance Matrices. *J. Mol. Biol.*, 233:123–138, 1993.

[108] M. J. Moutin, M. Cuillel, C. Rapin, R. Miras, M. Anger, A. M. Lompre, and Y. Dupont. Measurements of ATP binding on the large cytoplasmic loop of the sarcoplasmic reticulum Ca^{2+}-ATPase overexpressed in *Escherichia coli*. *J. Biol. Chem.*, 269:11147–11154, 1994.

[109] M. Bramkamp, M. Gaßel, and K. Altendorf. FITC Binding Site and p-Nitrophenyl Phosphatase Activity of the Kdp-ATPase of *Escherichia coli*. *Biochemistry*, 43:4559–4567, 2004.

[110] D. D. Boehr, A. R. Farley, G. D. Wright, and J. R. Cox. Analysis of the π-π stacking interactions between the aminoglycoside antibiotic kinase APH(3')-IIIa and its nucleotide ligands. *Chem. Biol.*, 9:1209–1217, 2002.

[111] J. P. Gallivan and D. A. Dougherty. A computational study of cation-π interactions vs salt bridges in aqueous media: Implications for protein engineering. *J. Am. Chem. Soc.*, 122:870–874, 2000.

[112] L. W. Hung, I. X. Wang, K. Nikaido, P. Q. Liu, G. F. Ames, and S. H. Kim. Crystal structure of the ATP-binding subunit of an ABC transporter. *Nature*, 396:703–707, 1998.

[113] G. Verdon, S. V. Albers, B. W. Dijkstra, A. J. M. Driessen, and A.-M. W. H. Thunissen. Crystal Structures of the ATPase Subunit of the Glucose ABC Transporter from *Sulfolobus solfatiaricus*: Nucleotide-free and Nucleotide-bound Conformations. *J. Mol. Biol.*, 45:2615–2623, 2003.

[114] O. Ramaen, S. Masscheleyn, F. Duffieux, O. Pamlard, M. Oberkampf, J.-Y. Lallemand, V. Stoven, and E. Jacquet. Biochemical characterization and NMR studies of the nucleotide-binding domain 1 of multidrug-resistance-associated protein 1: evidence for interaction between ATP and Trp[653]. *Biochem. J.*, 376:749–756, 2003.

[115] T. L.-M. Sørensen, J. V. Møller, and P. Nissen. Phosphor Transfer and Calcium Ion Occlusion in the Calcium Pump. *Science*, 304:1672–1675, 2004.

[116] C. Toyoshima and T. Mizutani. Crystal structure of the calcium pump with a bound ATP analogue. *Nature*, published online:30 June, 2004.

[117] C. Xu, W. J. Rice, W. He, and D. L. Stokes. A Structural Model for the Catalytic Cycle of Ca^{2+}-ATPase. *J. Mol. Biol.*, 316:201–211, 2002.

[118] G. A. Scarborough. Why we must move on from the E1E2 model for the reaction cycle of the P-type ATPases. *J. Bioenerg. Biomembr.*, 35:193–201, 2003.

[119] G. A. Scarborough. Rethinking the P-type ATPase problem. *Trends Biochem. Sci.*, 28:581–584, 2003.

[120] W. Kühlbrandt. Biology, Structure and Mechanism of P-type ATPases. *Nat. Rev.*, 5:282–295, 2004.

[121] J. Okkeri, L. Laakkonen, and T. Haltia. The nucleotide-binding domain of the Zn^{2+}-transporting P-type ATPase from *Escherichia coli* carries a glycine motif that may be involved in binding of ATP. *Biochem. J.*, 377:95–105, 2003.

[122] S. Lutsenko and J. H. Kaplan. Organization of P-Type ATPases - Significance of Structural Diversity. *J. Mol. Evol.*, 34:15607–156613, 1995.

[123] S. M. Albelda and C. A. Buck. Integrins and other cell adhesion molecules. *FASEB J.*, 4:2868–2880, 1990.

[124] R. O. Hynes. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell*, 69:11–25, 1992.

[125] S. Dedhar. Integrins and signal transduction. *Curr. Opin. Hematol.*, 6:37–43, 1999.

[126] E. F. Plow, T. A. Haas, L. Zhang, J. Loftus, and J. W. Smith. Ligand binding to integrins. *J. Biol. Chem.*, 275:21785–21788, 2000.

[127] T. A. Springer. Predicted and experimental structures of integrins and β-propellers. *Curr. Opin. Struct. Biol.*, 12:802–813, 2002.

[128] K.-E. Gottschalk and H. Kessler. The Structures of Integrins and Integrin – Ligand Complexes: Implications for Drug Design and Signal Transduction. *Angew. Chem. Int. Ed.*, 41:3767–3774, 2002.

[129] D. Cox, T. Aoki, J. Seki, Y. Motoyama, and K. Yoshida. The pharmacology of the integrins. *Med. Res. Rev.*, 14:195–228, 1994.

[130] K.-E. Gottschalk, P. D. Adams, A. T. Brunger, and H. Kessler. Transmembrane signal transduction of the $\alpha_{IIb}\beta_3$ integrin. *Protein Sci.*, 11:1800–1812, 2002.

[131] K.-E. Gottschalk, R. Günther, and H. Kessler. A Three-State Mechanism of Integrin Activation and Signal Transduction for Integrin $\alpha_V\beta_3$. *ChemBioChem*, 5:470–473, 2002.

[132] K.-E. Gottschalk and H. Kessler. Evidence for hetero-association of transmembrane helices of integrins. *FEBS Letters*, 557:252–258, 2004.

[133] O. Vinogradova, A. Velyvis, A. Velyviene, B. Hu, T. Haas, E. Plow, and J. Qin. A structural mechanism of integrin αIIbβ3 "inside-out" activation as regulated by its cytoplasmic face. *Cell*, 110:587–597, 2002.

[134] K.-E. Gottschalk and H. Kessler. A computational Model of Transmembrane Integrin Clustering. *Structure*, 12:1109–1116, 2004.

[135] T. A. Springer. Traffic signals for lymphocyte recirculation and leukocyte emigration: the multistep paradigm. *Cell*, 76:301–314, 1994.

[136] E. C. Butcher and L. J. Picker. Lymphocyte homing and homeostasis. *Science*, 272:60–66, 1996.

[137] N. Wagner, J. Lohler, E. J. K. und K Ley, E. Leung, G. Krissansen, K. Rajewsky, and W. Müller. Critical role for β7 integrins in formation of the gut-associated lymphoid tissue. *Nature*, 153:366–370, 1996.

[138] M. J. Briskin, L. Rott, and E. C. Butcher. Structural requirements for mucosal vascular addressin binding to its lymphocyte receptor α4β7. Common themes among integrin-Ig family interactions. *J. Immunol.*, 156:719–726, 1996.

[139] J. L. Viney, S. Jones, H. H. Chiu, B. Lagrimas, M. E. Renz, L. G. Presta, D. Jackson, K. J. Hillan, S. Lew, and S. Fong. Mucosal addressin cell

adhesion molecule-1: a structural and functional analysis demarcates the integrin binding motif. *J. Immunol.*, 157:2488–2497, 1996.

[140] J. Iwasa, T. Fujita, and C. Hansch. Subtituent constants for aliphatic functions obtained from partition coefficients. *J. Med. Chem.*, 56:150–153, 1965.

[141] C. A. Lipinski, F. Lombardo, B. W. Dominy, and P. J. Feeney. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug. Del. Rev.*, 23:3–25, 1997.

[142] C. A. Lipinski. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods*, 44:235–249, 2000.

[143] H. Kessler, H. R. Loosli, and H. Oschkinat. Assignment of the ^{1}H-, ^{13}C, and ^{15}N-NMR spectra of cyclosorin A in $CDCl_3$ and C_6D_6 by a combination of homo- and heteronuclear two-dimensional techniques. *Helv. Chim. Acta*, 68:661–682, 1985.

[144] H. R. Loosli, H. Kessler, H. Oschkinat, H. P. Weber, T. J. Petcher, and A. Widmer. The conformation of cyclosporin A in the crystal and in solution. *Helv. Chim. Acta*, 68:682–704, 1985.

[145] J. Lautz, H. Kessler, J. M. Blaney, R. M. Scheek, and W. F. van Gunsteren. Calculating three-dimensional molecular structure from atom-atom distance information: cyclosporin A. *J. Appl. Crystallog.*, 33:281–288, 1989.

[146] D. F. Veber, S. R. Johnson, H. Cheng, B. R. Smith, K. Ward, and K. D. Kopple. Molecular properties that influence the oral bioavailability of drug candidates. *J. Med. Chem.*, 330:343–358, 2002.

[147] J. Boer, D. Gottschling, A. Schuster, M. Semmrich, B. Holzmann, and H. Kessler. Design and Synthesis of Potent and Selective $\alpha 4\beta 7$ Integrin Antagonists. *J. Med. Chem.*, 44:2586–2592, 2001.

[148] D. Gottschling, J. Boer, L. Marinelli, G. Voll, M. Haupt, A. Schuster, B. Holzmann, and H. Kessler. Synthesis and NMR Structure of Peptidomimetic $\alpha 4\beta 7$ Integrin Antagonists. *ChemBioChem*, 3:575–578, 2002.

[149] H. Kessler. Peptoids - A New Approach to the Development of Pharmaceuticals. *Angew. Chem. Int. Ed. Engl.*, 32:543–544, 1993.

[150] E. K. Bradley. A Method for Sequential NMR Assignment of ^{1}H and [13]C Resonances of N-Substituted Glycine Peptoids. *J. Magn. Reson. B*, 110:195–197, 1996.

[151] H.-J. Lee, I.-A. Ahn, S. Ro, K.-H. Choi, Y.-S. Choi, and K.-B. Lee. Role of azaamino acid residue in beta-turn formation and stability in designed peptide. *J. Peptide Res.*, 56:35–46, 2000.

[152] F. André, G. Boussard, D. Bayeul, C. Didierjean, A. Aubry, and M. Marraud. Aza-peptides. II. X-ray structures of aza-alanine and aza-asparagine-containing peptides. *J. Peptide Res.*, 49:556–562, 1997.

[153] F. André, A. Vicherat, G. Boussard, A. Aubry, and M. Marraud. Aza-peptides. III. Experimental structural analysis of aza-alanine and aza-asparagine-containing peptides. *J. Peptide Res.*, 50:372–381, 1997.

[154] D. Gottschling. *Parallele Festphasensynthese und biologische Evaluation niedermolekularer Integrinantagonisten*. PhD thesis, Universität München, 2001.

[155] C. Palomo, J. M. Aizpurua, A. Benito, J. I. Miranda, R. M. Fratila, C. Matute, M. Domercq, F. Gago, S. Martin-Santamaria, and A. Linden. Development of a new family of conformationally restricted peptides as potent nucleators of β-turns. Design, synthesis, structure, and

biological evaluation of a β-lactam peptide analogue of melanostatin. *J. Am. Chem. Soc.*, 125:16243–16260, 2003.

[156] N. J. Baxter and M. P. Williamson. Temperature dependence of ^{1}H chemical shifts in proteins. *J. Biomol. NMR*, 9:359–369, 1997.

[157] T. F. Havel. *Distance Geometry and Molecular Conformation*. John Wiley & Sons, New York, 1988.

[158] A. T. Hagler. *Molecular Dynamics*, pages 133–138. Polycrystal Book Service, Western Springs, 1985.

[159] A. P. Nanzer, W. F. van Gunsteren, and A. E. Torda. Parametrisation of time-averaged distance restraints in MD simulations. *J. Biomol. NMR*, 6:313–320, 1995.

[160] M. Tiotto, V. Saudek, and V. Sklenar. Gradient-Tailored Excitation for Single-Quantum NMR Spectroscopy of Aqueous Solution. *J. Biomol. NMR*, 2:661–666, 1992.

Curriculum Vitae

Personal Data

Melina Haupt
born the 10$^{\text{th}}$ April 1975 in Traunstein, Germany

Education

1981–1985	Volksschule Chieming
1985–1994	Gymnasium Landschulheim Schloß Ising am Chiemsee
06/1994	Abitur
11/1994	Study of Chemistry, Ludwig-Maximilians Universität München
05/1997	Prediploma
11/98–02/99	Research project at the University of Glasgow, GB Section Inorganic Chemistry, Groups of Prof. Dr. J. M. Winfield and Dr. R. J. Cross *Synthesis of Trimethylplatiniumchloride and Phosphorligands*
01–07/2000	Diploma thesis at the Faculty of Organic Chemistry, LMU München, Group of Prof. Dr. W. Steglich *Isolierung und Strukturaufklärung des roten Pigments aus der Seeanemone* Urticina felina *und Studie zur Wechselwirkung des Antitumoragens Oxaliplatin mit Lactose*
07/2000	Diploma
since 11/2000	Doctoral thesis at the Chemistry Department, Technical University of Munich

Trainings and Jobs

08–10/1996	Trainee at SKW-Trostberg AG, Germany
08/1998	Work at SKW-Trostberg AG, Germany, R&D
08–10/2000	Trainee at SKW-Biosystems, Baupte, France, R&D

Publications

- *Trimethylplatinum hydroxy-fluorides*, R. J. Cross, M. Haupt, D. S. Rycroft, J. M. Winfield, *J. Organomet. Chem.*, **1999**, *578*, 195–199.

- *Synthesis and NMR Structure of Peptidomimetics as α4β7 Integrin Antagonists*, D. Gottschling, J. Boer, L. Marinelli, G. Voll, M. Haupt, A. Schuster, B. Holzmann and H. Kessler, *ChemBioChem*, **2002**, *3*, 575-578.

- *Letter to the Editor: ^1H, ^{13}C and ^{15}N resonance assignment of the nucleotide binding domain of KdpB from Escherichia coli*, M. Haupt, M. Coles, V. Truffault, M. Bramkamp, K. Altendorf and H. Kessler, *J. Biomol. NMR*, **2004**, *29*, 437–438.

- *Inter-domain motions of the N-domain of the KdpFABC complex, a P-type ATPase, are not driven by ATP-induced conformational changes*, M. Haupt, M. Bramkamp, M. Coles, K. Altendorf and H. Kessler, *J. Mol. Biol.*, **2004**, *342*, 1547–1558.

Contributions to International Conferences and Meetings

- Posterpresentation: *NMR Structure of Peptidomimetics as α4β7 Integrin Antagonists*, M. Haupt, D. Gottschling, J. Boer, L. Marinelli, G. Voll, H. Kessler, 16th European Experimental NMR Conference, Prague, Tchech Republic, 09.-14.06.2002.

- Oral Presentation: *KdpBN: Towards the solution NMR structure of the nucleotide binding domain of a potent K^+ uptake system*, M. Haupt, H. Kessler, M. Bramkamp, K. Altendorf, Protein Chemistry Ph.D. Summer School on Protein Folding and Protein Ligand Interactions, Spetses, Greece, Organisation: University of Copenhagen, Denmark, 20.–27.09.2002.

- Posterpresentation: *NMR Investigations on the Nucleotide Binding Domain of Kdp*, M. Haupt, M. Coles, H. Kessler, M. Bramkamp, K. Altendorf, 44th Experimental NMR Conference, Savannah, Georgia, USA, 30.03.–04.04.2003.

- Posterpresentation: *NMR Titration Studies on KdpBN*, M. Haupt, M. Coles, H. Kessler, M. Bramkamp, K. Altendorf, 25th Discussion Meeting, Leipzig, Germany, 30.09.–02.10.2003.

- Oral Presentation: *Insights in the Substrate Binding Mechanism of the Kdp-ATPase: Solution Structure of the Nucleotide Binding Domain of KdpB*, M. Haupt, M. Coles, H. Kessler, M. Bramkamp, K. Altendorf, DFG-Review Meeting, Priority Program 1070, Freiburg, Germany, 31.10.2003.

- Posterpresentation: *Insights in the Reaction Cycle of the Kdp System, a Unique P-type ATPase*, M. Haupt, M. Coles, H. Kessler, M. Bramkamp, K. Altendorf, 17th European Experimental NMR Conference, Lille, France, 06.–11.09.2004.

- Posterpresentation: *Inter-domain Motions of the N-domain of the KdpFABC Complex, a P-type ATPase, are not Driven by ATP-induced Conformational Changes*, M. Haupt, M. Coles, H. Kessler, M. Bramkamp, K. Altendorf, Annual Fall Meeting of the German Society for Biochemistry and Molecular Biology (GBM), Münster, Germany, 19.–22.09.2004.